

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

nucleic - nucleic search, using sw model

on: July 22, 2003, 11:16:54 ; Search time 985 Seconds  
(without alignments)  
1211.386 Million cell updates/sec

le: U95626-C-AT-42723  
fect score: 41  
uence: 1 tcagtctgagagcctga.....ataccaggactgcctgagac 41

ring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

rchd: 2054640 seqs, 14551402878 residues  
al number of hits satisfying chosen parameters: 4109280

imum DB seq length: 0  
imum DB seq length: 2000000000  
t-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

abase : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.rod.\*
- 36: em.htg.mam.\*
- 37: em.htg.vrt.\*
- 38: em.sy.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	41	100.0	583	9	HSMPA01
2	41	100.0	2082	9	AF068265
3	41	100.0	143068	6	AX335952
4	41	100.0	143068	9	HSU95626
5	41	100.0	185437	9	AC098613
6	27.2	66.3	120793	2	AF322450
7	27.2	66.3	133544	9	AC100748
8	27.2	66.3	163035	2	AC006405
9	25.6	62.4	110394	9	AC003086
10	25.6	62.4	142082	2	AC093137
11	25.6	62.4	159365	9	AL392111
12	25.6	62.4	174133	9	AC093183
13	25.6	62.4	176875	9	AC018638
14	25.6	62.4	178451	9	AL139396
15	25.6	62.4	179110	9	AC011755
16	25.6	62.4	179503	9	AL445468
17	25.6	62.4	183434	2	AC103924
18	25.6	62.4	180948	2	AC093138
19	25.6	62.4	210736	2	AC092792
20	25.2	61.5	132463	2	AC111775
21	25.2	61.5	187803	10	AL596252
22	25.2	61.5	188780	9	AC010896
23	25	61.0	68116	2	AC103970
24	25	61.0	73235	2	AC103725
25	25	61.0	78986	2	AC115626
26	25	61.0	82601	9	AL136971
27	25	61.0	112204	9	AL136439
28	25	61.0	116845	9	AL133174
29	25	61.0	116845	9	AC008819
30	25	61.0	117583	9	AC034238
31	25	61.0	120689	9	CNS01RGS
32	25	61.0	133322	2	AC091845
33	25	61.0	142441	2	HSJ344H20
34	25	61.0	155428	9	AC009682
35	25	61.0	160979	2	AC055879
36	25	61.0	164809	9	AC128690
37	25	61.0	165146	2	AC011644
38	25	61.0	167163	2	AC013705
39	25	61.0	169337	2	AL356462
40	25	61.0	173834	9	AC108044
41	25	61.0	174701	9	AC007625
42	25	61.0	182019	2	AC024388
43	25	61.0	187889	9	AC093008
44	25	61.0	193131	9	AL390766
45	25	61.0	202750	2	AC116942

ALIGNMENTS

RESULT 1	HSMPA01	583 bp	DNA	linear	PRI 24-JUL-1997
LOCUS	Human monocyte chemoattractant protein 1				receptor gene, 5' region.
DEFINITION	U80923				
ACCESSION	U80923.1	GI:1773031			
VERSION	1				
KEYWORDS	1 of 2				
SEGMENT	Homo sapiens				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 583)				
AUTHORS	Wong, L.M., Myers, S.J., Tsou, C.L., Gosling, J., Arai, H. and Charo, I.F.				

TITLE Organization and differential expression of the human monocyte chemoattractant protein 1 receptor gene. Evidence for the role of the carboxyl-terminal tail in receptor trafficking

J. Biol. Chem. 272 (2), 1038-1045 (1997)

79150864

8995400

2 (bases 1 to 583)

Myers, S.J. and Charo, I.F.

Direct Submission

Submitted (04-DEC-1996) Pharmacology, Emory University, 1510

Clifton Road, Atlanta 30322, USA

Location/Qualifiers

1..583

/organism="Homo sapiens"

/db\_xref="taxon:9606"

157 a 136 c 156 g 134 t

Query Match 100.0%; Score 41; DB 9; Length 583;

Best Local Similarity 100.0%; Pred. No. 5.6e-05;

Mismatches 0; Indels 0; Gaps 0;

1 TCAGTGTCTGAGAGCTGACATACACAGAGCTGCTGAGAC 41

|||||

62 TCAGTGTCTGAGAGCTGACATACACAGAGCTGCTGAGAC 102

UT 2

68265

US

INITIATION

gene promoter and mRNA, partial sequence.

AF068265

AF068265

AF068265.1 GI:4587865

WORDS

ACE

ORGANISM

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 2082)

Yamamoto, K., Takeshima, H., Hamada, K., Nakao, M., Kino, T., Nishi, T.,

Kochi, M., Kuratsu, J., Yoshimura, T. and Ushio, Y.

Cloning and functional characterization of the 5'-flanking region

of the human monocyte chemoattractant protein-1 receptor (CCR2)

gene. Essential role of 5'-untranslated region in tissue-specific

expression

J. Biol. Chem. 274 (8), 4646-4654 (1999)

99143121

998701

2 (bases 1 to 2082)

Yamamoto, K., Takeshima, H., Hamada, K., Nakao, M., Kino, T., Nishi, T.,

Kochi, M., Kuratsu, J., Yoshimura, T. and Ushio, Y.

Direct Submission

Submitted (26-MAY-1998) Neurosurgery, Kumamoto University School of

Medicine, Honjo, 1-1-1, Kumamoto 860-8556, Japan

Location/Qualifiers

1..2082

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/map="3p"

1..>2082

/gene="CCR2"

1..1697

/gene="CCR2"

1570..1574

/gene="CCR2"

1666..1672

/gene="CCR2"

1998..>2012

/gene="CCR2"

/product="monocyte chemoattractant protein 1 receptor"

1698..2012

5'UTR

BASE COUNT 650 a 439 c 433 g 560 t

ORIGIN

Query Match 100.0%; Score 41; DB 9; Length 2082;

Best Local Similarity 100.0%; Pred. No. 4.3e-05;

Mismatches 0; Indels 0; Gaps 0;

1 TCAGTGTCTGAGAGCTGACATACACAGAGCTGCTGAGAC 41

|||||

1959 TCAGTGTCTGAGAGCTGACATACACAGAGCTGCTGAGAC 1999

QY

Db

RESULT 3

AX335952

LOCUS

Sequence 6461 from Patent WO0194629.

AX335952

AX335952.1 GI:18126671

KEYWORDS

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1

Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,

Horrigan, S., Soppet, D.R. and Weaver, Z.

Cancer gene determination and therapeutic screening using signature

gene sets

Patent: WO 0194629-A 6461 13-DEC-2001;

Avalon Pharmaceuticals (US)

Location/Qualifiers

1..143068

/organism="Homo sapiens"

/db\_xref="taxon:9606"

41194 a 30122 c 32403 g 39349 t

BASE COUNT

ORIGIN

Query Match 100.0%; Score 41; DB 6; Length 143068;

Best Local Similarity 100.0%; Pred. No. 1.9e-05;

Mismatches 0; Indels 0; Gaps 0;

1 TCAGTGTCTGAGAGCTGACATACACAGAGCTGCTGAGAC 41

|||||

42703 TCAGTGTCTGAGAGCTGACATACACAGAGCTGCTGAGAC 42743

QY

Db

RESULT 4

HSU95626

LOCUS

Sequence 143068 bp DNA linear PRI 16-MAY-1997

DEFINITION

Homo sapiens ccr2b (ccr2), ccr2a (ccr2), ccr5 (ccr5) and ccr6

(ccr6) genes, complete cds, and lactoferrin (lactoferrin) gene,

partial cds, complete sequence.

U95626

U95626.1 GI:2104517

ACCESSION

U95626

VERSION

U95626

KEYWORDS

HTG.

SOURCE

Homo sapiens.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 143068)

McCombie, R.W., Wilson, R., Chen, E., Gibbs, R., Zuo, L., Johnson, D.,

Nhan, M., Parnell, L., Dedhia, N., Ansari, A., Mardis, E., Schutz, K.,

Gao, J., de la Bastide, M., Kaplan, N., Greco, T., Touchman, J.,

Muzny, D., Chen, C.-N., Evans, C., FitzGerald, M., See, L.H., Tang, M.,

Porcel, B.M., Dragan, Y., Giacalone, J., Pae, A., Powell, E.,

Solinsky, K.A., Desilva, U., Diaz-Perez, S., Zhou, X., Yu, Y.,

Watanabe, M., Doggett, N., Garcia, D. and Sagripanti, J.-L.

Human BAC clone 110P12

Unpublished (1997)

2 (bases 1 to 143068)

McCombie, R.W., Wilson, R., Chen, E., Gibbs, R., Zuo, L., Johnson, D.,

Nhan, M., Parnell, L., Dedhia, N., Ansari, A., Mardis, E., Schutz, K.,

Gao, J., de la Bastide, M., Kaplan, N., Greco, T., Touchman, J.,

Muzny, D., Chen, C.-N., Evans, C., FitzGerald, M., See, L.H., Tang, M.,

Porcel, B.M., Dragan, Y., Giacalone, J., Pae, A., Powell, E.,

Solinsky, K.A., Desilva, U., Diaz-Perez, S., Zhou, X., Yu, Y.,

Watanabe, M., Doggett, N., Garcia, D. and Sagripanti, J.-L.

Human BAC clone 110P12

Unpublished (1997)

2 (bases 1 to 143068)

McCombie, R.W., Wilson, R., Chen, E., Gibbs, R., Zuo, L., Johnson, D.,

Nhan, M., Parnell, L., Dedhia, N., Ansari, A., Mardis, E., Schutz, K.,

Gao, J., de la Bastide, M., Kaplan, N., Greco, T., Touchman, J.,

Muzny, D., Chen, C.-N., Evans, C., FitzGerald, M., See, L.H., Tang, M.,

Porcel, B.M., Dragan, Y., Giacalone, J., Pae, A., Powell, E.,

Solinsky, K.A., Desilva, U., Diaz-Perez, S., Zhou, X., Yu, Y.,

Watanabe, M., Doggett, N., Garcia, D. and Sagripanti, J.-L.

Human BAC clone 110P12

Unpublished (1997)

2 (bases 1 to 143068)

McCombie, R.W., Wilson, R., Chen, E., Gibbs, R., Zuo, L., Johnson, D.,

Nhan, M., Parnell, L., Dedhia, N., Ansari, A., Mardis, E., Schutz, K.,

Gao, J., de la Bastide, M., Kaplan, N., Greco, T., Touchman, J.,

Muzny, D., Chen, C.-N., Evans, C., FitzGerald, M., See, L.H., Tang, M.,

Porcel, B.M., Dragan, Y., Giacalone, J., Pae, A., Powell, E.,

Solinsky, K.A., Desilva, U., Diaz-Perez, S., Zhou, X., Yu, Y.,

Watanabe, M., Doggett, N., Garcia, D. and Sagripanti, J.-L.

Human BAC clone 110P12

Unpublished (1997)

2 (bases 1 to 143068)

McCombie, R.W., Wilson, R., Chen, E., Gibbs, R., Zuo, L., Johnson, D.,

Nhan, M., Parnell, L., Dedhia, N., Ansari, A., Mardis, E., Schutz, K.,

Gao, J., de la Bastide, M., Kaplan, N., Greco, T., Touchman, J.,

Muzny, D., Chen, C.-N., Evans, C., FitzGerald, M., See, L.H., Tang, M.,

Porcel, B.M., Dragan, Y., Giacalone, J., Pae, A., Powell, E.,

Solinsky, K.A., Desilva, U., Diaz-Perez, S., Zhou, X., Yu, Y.,

Watanabe, M., Doggett, N., Garcia, D. and Sagripanti, J.-L.

Human BAC clone 110P12

Unpublished (1997)

2 (bases 1 to 143068)

McCombie, R.W., Wilson, R., Chen, E., Gibbs, R., Zuo, L., Johnson, D.,

Nhan, M., Parnell, L., Dedhia, N., Ansari, A., Mardis, E., Schutz, K.,

Gao, J., de la Bastide, M., Kaplan, N., Greco, T., Touchman, J.,

Muzny, D., Chen, C.-N., Evans, C., FitzGerald, M., See, L.H., Tang, M.,

Porcel, B.M., Dragan, Y., Giacalone, J., Pae, A., Powell, E.,

Solinsky, K.A., Desilva, U., Diaz-Perez, S., Zhou, X., Yu, Y.,

Watanabe, M., Doggett, N., Garcia, D. and Sagripanti, J.-L.

Human BAC clone 110P12

Unpublished (1997)

2 (bases 1 to 143068)

McCombie, R.W., Wilson, R., Chen, E., Gibbs, R., Zuo, L., Johnson, D.,

Nhan, M., Parnell, L., Dedhia, N., Ansari, A., Mardis, E., Schutz, K.,

Gao, J., de la Bastide, M., Kaplan, N., Greco, T., Touchman, J.,

Muzny, D., Chen, C.-N., Evans, C., FitzGerald, M., See, L.H., Tang, M.,

Porcel, B.M., Dragan, Y., Giacalone, J., Pae, A., Powell, E.,

Solinsky, K.A., Desilva, U., Diaz-Perez, S., Zhou, X., Yu, Y.,

Watanabe, M., Doggett, N., Garcia, D. and Sagripanti, J.-L.

Human BAC clone 110P12

Unpublished (1997)

2 (bases 1 to 143068)

McCombie, R.W., Wilson, R., Chen, E., Gibbs, R., Zuo, L., Johnson, D.,

Nhan, M., Parnell, L., Dedhia, N., Ansari, A., Mardis, E., Schutz, K.,

Gao, J., de la Bastide, M., Kaplan, N., Greco, T., Touchman, J.,

Muzny, D., Chen, C.-N., Evans, C., FitzGerald, M., See, L.H., Tang, M.,

Porcel, B.M., Dragan, Y., Giacalone, J., Pae, A., Powell, E.,

Solinsky, K.A., Desilva, U., Diaz-Perez, S., Zhou, X., Yu, Y.,

Watanabe, M., Doggett, N., Garcia, D. and Sagripanti, J.-L.

Human BAC clone 110P12

Unpublished (1997)

2 (bases 1 to 143068)

McCombie, R.W., Wilson, R., Chen, E., Gibbs, R., Zuo, L., Johnson, D.,

Nhan, M., Parnell, L., Dedhia, N., Ansari, A., Mardis, E., Schutz, K.,

Gao, J., de la Bastide, M., Kaplan, N., Greco, T., Touchman, J.,

Muzny, D., Chen, C.-N., Evans, C., FitzGerald, M., See, L.H., Tang, M.,

Porcel, B.M., Dragan, Y., Giacalone, J., Pae, A., Powell, E.,

Solinsky, K.A., Desilva, U., Diaz-Perez, S., Zhou, X., Yu, Y.,

Watanabe, M., Doggett, N., Garcia, D. and Sagripanti, J.-L.

Human BAC clone 110P12

Unpublished (1997)

2 (bases 1 to 143068)

McCombie, R.W., Wilson, R., Chen, E., Gibbs, R., Zuo, L., Johnson, D.,

Nhan, M., Parnell, L., Dedhia, N., Ansari, A., Mardis, E., Schutz, K.,

Gao, J., de la Bastide, M., Kaplan, N., Greco, T., Touchman, J.,

Muzny, D., Chen, C.-N., Evans, C., FitzGerald, M., See, L.H., Tang, M.,

Porcel, B.M

Gnoj, L., de la Bastide, M., Kaplan, N., Greco, T., Touchman, J., Muzny, D., Chen, C.-N., Evans, C., FitzGerald, M., See, L. H., Tang, M., Porcel, B. M., Dragan, Y., Giacalone, J., Pae, A., Powell, E., Solinsky, K. A., DeSilva, U., Diaz-Perez, S., Zhou, X., Yu, Y., Watanabe, M., Doggett, N., Garcia, D. and Sagripanti, J.-L.  
 Direct Submission  
 Submitted (27-MAR-1997) Advanced Genome Sequence Analysis Course, Cold Spring Harbor Laboratory, 1Bungtown Rd., Cold Spring Harbor, NY 11724, USA  
 Regions with single-strand coverage are as follows:

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

# TURES

## source

1. 143068  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="3"  
 /clone="BAC 110P12"  
 46056. .49505  
 /gene="ccr2"  
 /note="Confirmed by similarity to Human monocytic chemoattractant protein 1 receptor (ccr2) mRNA (Accession Number U80924), two alternatively spliced mRNAs."  
 Join(46056. .47046,48255. .49505)  
 /gene="ccr2"  
 /product="ccr2a"  
 /note="Confirmed by similarity to Human monocytic chemoattractant protein 1 receptor (ccr2) alternatively spliced mRNA encoding A-form carboxyl tail, Accession Number U80924."  
 46056. .47997  
 /gene="ccr2"  
 /product="ccr2b"  
 /note="Confirmed by similarity to Human monocytic chemoattractant protein 1 receptor (ccr2) alternatively spliced mRNA encoding B-form carboxyl tail. Accession Number: U80924."  
 Join(46106. .47046,48255. .48438)  
 /gene="ccr2"  
 /note="Confirmed by similarity to Human monocytic chemoattractant protein 1 receptor (ccr2) alternatively spliced A-form, Encoded by GenBank Accession Number U80924, gi 1168965"  
 /codon\_start=1  
 /product="ccr2a"  
 /db\_xref="GI:2104518"  
 /translation="MLSTSRSPRIETNESGEVITFFDYDYGAPCHKFDVQKGAQL LPPLYSIVFIPIGFVGNMLVLLINCKKLCITDIYLLINLAISDLFLITPLMAHSA ANEWFGNMCKFTGLTGHVFGGFIITLDIYLAIVHAFALKARTVFGVVT SVITWLVAASVPGIIFTKCKEDSVYVCGPFPGRWNFMHIMNLIGLVPLIM VICYSGILKTLRCNKKHRAVRVITMIVFLFWPTYNIVLLNTFOFFGLSN CESTSOLDQAQVETETGWHCTCINPIIYAFVGEKFRYLSVFRKXHTKRFCKQCPV

## gene

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

## mRNA

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

## mRNA

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

## CDS

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

## CDS

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

## gene

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

## mRNA

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

## CDS

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

# TURES

## source

1. 143068  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="3"  
 /clone="BAC 110P12"  
 46056. .49505  
 /gene="ccr2"  
 /note="Confirmed by similarity to Human monocytic chemoattractant protein 1 receptor (ccr2) mRNA (Accession Number U80924), two alternatively spliced mRNAs."  
 Join(46056. .47046,48255. .49505)  
 /gene="ccr2"  
 /product="ccr2a"  
 /note="Confirmed by similarity to Human monocytic chemoattractant protein 1 receptor (ccr2) alternatively spliced mRNA encoding A-form carboxyl tail, Accession Number U80924."  
 46056. .47997  
 /gene="ccr2"  
 /product="ccr2b"  
 /note="Confirmed by similarity to Human monocytic chemoattractant protein 1 receptor (ccr2) alternatively spliced mRNA encoding B-form carboxyl tail. Accession Number: U80924."  
 Join(46106. .47046,48255. .48438)  
 /gene="ccr2"  
 /note="Confirmed by similarity to Human monocytic chemoattractant protein 1 receptor (ccr2) alternatively spliced A-form, Encoded by GenBank Accession Number U80924, gi 1168965"  
 /codon\_start=1  
 /product="ccr2a"  
 /db\_xref="GI:2104518"  
 /translation="MLSTSRSPRIETNESGEVITFFDYDYGAPCHKFDVQKGAQL LPPLYSIVFIPIGFVGNMLVLLINCKKLCITDIYLLINLAISDLFLITPLMAHSA ANEWFGNMCKFTGLTGHVFGGFIITLDIYLAIVHAFALKARTVFGVVT SVITWLVAASVPGIIFTKCKEDSVYVCGPFPGRWNFMHIMNLIGLVPLIM VICYSGILKTLRCNKKHRAVRVITMIVFLFWPTYNIVLLNTFOFFGLSN CESTSOLDQAQVETETGWHCTCINPIIYAFVGEKFRYLSVFRKXHTKRFCKQCPV

## gene

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

## mRNA

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

## CDS

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

## CDS

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 11

CAPSPSPRNLCALCIGDEQENKCVPSNERYGTTGAFRCLEAGADVAFKDVTV  
LONTGNNNDWAKDLKALFALLCLDGRKXVTEARSCHLAPNPAHVSRMDKVER  
LKQVLLHQAKFGRNGSDPCDFCLFQSTKMLFNDFNTECLARLKGKTYIKILGPQ  
YVAGITNLKCKSTSPLEACEFLRK"

3E COUNT 41194 a 30122 c 32403 g 39349 t  
GIN

Query Match 100.0%; Score 41; DB 9; Length 143068;  
Best Local Similarity 100.0%; Pred. No. 1.9e-05;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TCAGTGTCTGAGAGCCTGACATACCAAGACTGCCTGAGAC 41

42703 TCAGTGTCTGAGAGCCTGACATACCAAGACTGCCTGAGAC 42743

AC098613 185437 bp DNA linear PRI 01-AUG-2002  
AC098613 Homo sapiens chromosome 3 clone RP11-24F11, complete sequence.  
AC098613.2 GI:22038607  
HTG.  
JRCF human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 185437)  
Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,  
Saenphimmachak,C., Phelps,K.A., Buckley,D., Kibukawa,M., Raymond,C.  
and Haugen,E.D.  
Direct Submission  
Unpublished  
2 (bases 1 to 185437)  
Kaul,R.K., Olson,M.V., Raymond,C. and Haugen,E.D.  
Direct Submission  
Submitted (26-OCT-2001) Genome Center, University of Washington,  
Box 352145, Seattle, WA 98195, USA  
3 (bases 1 to 185437)  
Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,  
Saenphimmachak,C., Phelps,K.A., Buckley,D., Kibukawa,M., Raymond,C.  
and Haugen,E.D.  
Direct Submission  
Submitted (01-AUG-2002) Genome Center, University of Washington,  
Box 352145, Seattle, WA 98195, USA  
On Aug 1, 2002 this sequence version replaced gi:16445164.

Center: University of Washington Genome Center  
Center Code: UWGC  
Web site: <http://www.genome.washington.edu>  
Contact: [uwgctgs@u.washington.edu](mailto:uwgctgs@u.washington.edu)  
Project Information  
Center project name: chr-3  
Center clone name: RP11-24F11 (bc0137)  
Summary Statistics  
Sequencing vector: plasmid; 100% of reads  
Chemistry: Dye-terminator ET; 93% of reads  
Chemistry: Dye-terminator Big Dye; 7% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 184860 bases at least Q40  
Consensus quality: 185398 bases at least Q30  
Consensus quality: 185435 bases at least Q20  
Insert size: 185437; sum-of-contigs  
Quality coverage: 7.6x in Q20 bases; sum-of-contigs

Overlapping Sequences:  
5': BAC-110P12 U95626, 111014-bp overlap  
3': RP11-509I21 (UWGC:bc0454) AC104304, 61294-bp overlap  
Sequence Quality Assessment:  
This entry has been annotated with sequence quality  
estimates computed by the Phrap assembly program.  
All manually edited bases have been reduced to quality zero.

Quality levels above 40 are expected to have less than  
1 error in 10,000 bp.  
Base-by-base quality values are not generally visible from the  
GenBank flat file format but are available as part  
of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted:  
all regions were either double-stranded or sequenced with an  
alternate chemistry or covered by high quality data (i.e., Phred  
quality >= 30); an attempt was made to resolve all sequencing  
problems, such as compressions and repeats; all regions were  
covered by at least one plasmid subclone or more than one M13  
subclone; and the assembly was confirmed by restriction digest.

#### Sequence Validation:

This sequence has been validated by Multiple Complete Digest  
fingerprinting. Comparison of the experimentally derived digest  
fragments with sequence-predicted fragments is given below.  
The electronically-digested sequence consists of both insert and  
vector in order to accurately represent the entire circular BAC.  
Small fragments below a variable cutoff (approximately 400-800 bp)  
are not resolved in the fingerprint and hence do not appear  
in the table. There are no significant remaining discrepancies  
between the experimental and predicted values. Uniquely ordered  
fragments are separated by dashed lines.

BglII				EcoRI				HindIII			
SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt
11786	12040	8696	8837	8837	9531	9423					
2067	2065	6	<800	<800	6382	6501					
5681	5720	1846	1824	1824	512	<800					
3716	3953	4052	3986	3986	449	<800					
875	897	1159	1130	1130	6692	6501					
910	897	54	<800	<800	10449	10229					
2169	2215	560	<800	<800	953	1025					
1705	1653	2287	2267	2267	3124	3176					
5763	5720	4905	4891	4891	1054	1025					
5844	5720	3049	3078	3078	3985	3998					
5061	5001	1377	1394	1394	1100	1025					
2625	2640	9903	9772	9772	124	<800					
725	<800	1022	1021	1021	2671	2693					
1173	1161	12606	12503	12503	1948	1968					
5747	5720	866	868	868	83	<800					
9995	9684	9817	9772	9772	7455	7474					
999	999	3598	3601	3601	1305	1267					
4541	4503	452	<800	<800	1047	1025					
2950	2981	7549	7635	7635	3299	3301					
406	<800	2063	2075	2075	5279	5237					
3375	3602	5837	5858	5858	6815	6926					

13024	13045	13685	13328	435	<800	Query Match
416	<800	4104	3986	2509	2576	Best Local Similarity 100.0%; Score 41; DB 9; Length 185437;
10298	10140	1943	1928	2218	2172	Mismatches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1478	1452	3964	3986	3797	3802	
287	<800	2758	2771	7758	7830	QY 1 TCAGTTGCTGAGAGCTGACATACCAGGACTGCCTGAGAC 41
205	<800	167	<800	16192	16442	DB 10653 TCAGTTGCTGAGAGCTGACATACCAGGACTGCCTGAGAC 10693
9428	9684	1307	1304	11020	10820	
3809	3953	13306	13328	1978	1968	RESULT 6
1275	1234	2423	2450	8741	8681	AF322450
3351	3381	2243	2267	4795	4815	LOCUS AF322450
193	<800	1109	1130	14921	14985	DEFINITION Homo sapiens chromosome 17 clone BAC407121 map 17p13.3, ***
334	<800	4131	3986	1896	1968	SEQUENCING IN PROGRESS ***, 28 unordered pieces.
2546	2640	6383	6450	172	<800	ACCESSION AF322450.1 GI:11559855
9155	9107	332	<800	2091	2172	VERSION AF322450
9190	9107	242	<800	2159	2172	KEYWORDS HTG; HTGS PHASE1.
352	<800	15571	15803	3573	3564	SOURCE Homo sapiens
1157	1161	453	<800	5252	5237	ORGANISM Homo sapiens
3967	3953	1429	1394	4478	4428	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1369	1334	655	<800	6166	6243	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
752	752	7881	7987	1565	1603	REFERENCE 1 (bases 1 to 120793)
1624	1653	7533	7635	1645	1603	Zhao,X.T., He,M., Wan,D.F., Ye,Y., Qin,W.X., Huang,Y., Zuo,L. and
3258	3381	585	<800	573	<800	Gu,J.R.
9624	9684	2998	2973	7986	7830	TITLE Gene clone on human chromosome 17p13.3
2841	2870	8796	8837	2313	2324	JOURNAL Unpublished
2979	3069	1331	1304	49	<800	REFERENCE 2 (bases 1 to 120793)
3898	3953	7102	7161	1250	1267	Zhao,X.T., He,M., Wan,D.F., Ye,Y., Qin,W.X., Huang,Y., Zuo,L. and
4830	4804	---	---	2290	2324	Gu,J.R.
2418	2459	---	---	2056	1968	Direct Submission
6198	6066	---	---	---	---	Submitted (21-NOV-2000) National Laboratory For Oncogenes & Related
534	<800	---	---	---	---	Genes, Shanghai Cancer Institute, 25/Ln 2200 Xie-Tu Road, Shanghai
3985	3953	---	---	---	---	200032, P. R. China
1184	1161	---	---	---	---	* NOTE: This is a 'working draft' sequence. It currently
63	<800	---	---	---	---	* consists of 28 contigs. The true order of the pieces
						* is not known and their order in this sequence record is
						* arbitrary. Gaps between the contigs are represented as
						* runs of N, but the exact sizes of the gaps are unknown.
						* This record will be updated with the finished sequence
						* as soon as it is available and the accession number will
						* be preserved.
						* 1 608: contig of 608 bp in length
						* gap of unknown length
						* 609 1282: contig of 674 bp in length
						* gap of unknown length
						* 1283 1733: contig of 451 bp in length
						* gap of unknown length
						* 1734 2435: contig of 702 bp in length
						* gap of unknown length
						* 2436 2895: contig of 460 bp in length
						* gap of unknown length
						* 2896 3643: contig of 748 bp in length
						* gap of unknown length
						* 3644 4149: contig of 506 bp in length
						* gap of unknown length
						* 4150 4908: contig of 759 bp in length
						* gap of unknown length
						* 4909 5741: contig of 833 bp in length
						* gap of unknown length
						* 5742 6777: contig of 1036 bp in length
						* gap of unknown length
						* 6778 7766: contig of 989 bp in length
						* gap of unknown length
						* 7767 9183: contig of 1417 bp in length
						* gap of unknown length
						* 9184 9931: contig of 748 bp in length
						* gap of unknown length
						* 9932 12274: contig of 2343 bp in length
						* gap of unknown length
						* 12275 16059: contig of 3785 bp in length
						* gap of unknown length
						* 16060 19969: contig of 3910 bp in length
						* gap of unknown length

TURES  
source  
1. .185437  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="3"  
/clone="RP11-24F11"











JOURNAL Submitted (11-AUG-2001) NIH Intramural Sequencing Center, 8717  
 Government Circle, Gaithersburg, MD 20877, USA  
 REFERENCE 3 (bases 1 to 142082)  
 AUTHORS Green, E.D.  
 TITLE Direct Submission  
 JOURNAL Submitted (12-JUN-2002) NIH Intramural Sequencing Center, 8717  
 Government Circle, Gaithersburg, MD 20877, USA  
 COMMENT On Jun 12, 2002 this sequence version replaced gi:15148127.  
 ----- Genome Center  
 Center: NIH Intramural Sequencing Center  
 Center code: NISC  
 Web site: <http://www.nisc.nih.gov>  
 Contact: [nisc.zoo@nhgri.nih.gov](mailto:nisc.zoo@nhgri.nih.gov)  
 ----- Project Information  
 Center project name: anl  
 Center clone name: 144H14

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics  
 Sequencing vector: plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 141565 bases at least Q40  
 Consensus quality: 141698 bases at least Q30  
 Consensus quality: 141749 bases at least Q20  
 Insert size: 113000; agarose-fp  
 Insert size: 100000; pulse-field-gel  
 Insert size: 141782; sum-of-contigs  
 Quality coverage: 13.44x in Q20 bases; agarose-fp  
 Quality coverage: 15.18x in Q20 bases; pulse-field-gel  
 Quality coverage: 10.71x in Q20 bases; sum-of-contigs

-----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 4 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submittor.

\* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 \* 1 47102: contig of 47102 bp in length  
 \* 47103 47202: gap of unknown length  
 \* 47203 86175: contig of 38973 bp in length  
 \* 86176 86275: gap of unknown length  
 \* 86276 118675: contig of 32400 bp in length  
 \* 118676 118775: gap of unknown length  
 \* 118776 142082: contig of 23307 bp in length.

FEATURES  
 source  
 1. .142082  
 /organism="Pan troglodytes"  
 /db\_xref="taxon:9598"  
 /clone="RP43-144H14"  
 /clone\_lib="RP43"

misc\_feature  
 1. .47102  
 /note="assembly fragment  
 clone\_end:r7  
 vector\_side:left"  
 12236..142082

misc\_feature  
 /note="Clone overlaps with GenBank Accession Number  
 AC093138 clone RP43-154A20 (center project name anl)"  
 47203..86175

misc\_feature  
 86276..118675  
 /note="assembly fragment"  
 misc\_feature  
 118776..142082  
 /note="assembly fragment  
 clone\_end:SP6  
 vector\_side:right"  
 misc\_feature  
 132713..142082  
 /note="clone overlaps with GenBank Accession Number  
 AC093138 clone RP43-17P3 (center project name anl)"  
 BASE COUNT 42473 a 27490 c 27247 g 44571 t 301 others  
 ORIGIN

Query Match 62.4%; Score 25.6; DB 2; Length 142082;  
 Best Local Similarity 77.5%; Pred. No. 12;  
 Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 CAGTTGCTGAGAGCCTGACATACCGAGCTGCTGAGAC 41  
 |||||  
 Db 38952 CAGCTCTCAGAGAGCTGAGGTAGGAGCTGCTGAGCC 38991  
 |||||

RESULT 11  
 AL392111/c  
 LOCUS AL392111 159365 bp DNA linear PRI 15-NOV-2001  
 DEFINITION Human DNA sequence from clone RP11-48715 on chromosome 10, complete  
 sequence.  
 ACCESSION AL392111  
 VERSION AL392111.12 GI:16972970  
 KEYWORDS HTG.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 159365)  
 AUTHORS Bray-Allen, S.  
 TITLE Direct Submission  
 JOURNAL Submitted (15-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
[humquerry@sanger.ac.uk](mailto:humquerry@sanger.ac.uk) Clone requests: [clonerequest@sanger.ac.uk](mailto:clonerequest@sanger.ac.uk)  
 On Nov 16, 2001 this sequence version replaced gi:15808188.

During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. The following  
 abbreviations are used to associate primary accession numbers given  
 in the feature table with their source databases: Em: EMBL; Swi:  
 SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP  
 database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence  
 was generated from part of bacterial clone contigs of human  
 chromosome 10, constructed by the Sanger Centre Chromosome 10  
 Mapping Group. Further information can be found at  
<http://www.sanger.ac.uk/HGP/Chr10>  
 RP11-48715 is from the library RP11-11.2 constructed by the group  
 of Pieter de Jong. For further details see  
<http://www.chori.org/bacpac/home.htm>  
 VECTOR: pBACe3.6

IMPORTANT: This sequence is not the entire insert of clone  
 RP11-48715. It may be shorter because we sequence overlapping  
 sections only once, except for a short overlap.  
 The true right end of clone RP11-48715 is at 159365 in this  
 sequence. The true left end of clone RP11-375G3 is at 149681 in  
 this sequence. The true right end of clone RP11-77G23 is at 4000 in

```

this sequence.
Location/Qualifiers
1. .159365
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-48715"
/clone.lib="RP11-11.2"
103353 41838 a 35376 c 36999 g 45152 t
GIN

Query Match 62.4%; Score 25.6; DB 9; Length 159365;
est Local Similarity 77.5%; Pred. No. 12;
atches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

2 CAGTGTGCTGAGAGCTGACATACAGGATGCTGAGAC 41
|||||
103353 CAGCTACTCAGAGGCTGATGCGAGGAGGACTGCTGAGAC 103314

UT 12
93183/c
US
AC093183 174133 bp DNA linear PRI 29-MAR-2002
INITIATION Homo sapiens chromosome 7 clone RP11-78611, complete sequence.
SSION AC093183
STON AC093183.3 GI:19807864
WORDS HYG.
RCR Homo sapiens.
RGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 174133)
Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
Saenphimmachak,C., Phelps,K.A., Raymond,C. and Haugen,E.D.
Direct Submission
2 (bases 1 to 174133)
Kaul,R.K., Olson,M.V., Raymond,C., Clendenning,J., Ivey,R.G. and
Haugen,E.D.
Direct Submission
Submitted (14-AUG-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
3 (bases 1 to 174133)
Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
Saenphimmachak,C., Phelps,K.A., Raymond,C. and Haugen,E.D.
Direct Submission
Submitted (29-MAR-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
On Mar 29, 2002 this sequence version replaced gi.17998609.
----- Genome Center
Center: University of Washington Genome Center
Center Code: UWGC
Web site: http://www.genome.washington.edu
Contact: uwgctgs@u.washington.edu
----- Project Information
Center project name: chr-7
Center clone name: RP11-78611 (djs736)
----- Summary Statistics
Sequencing vector: plasmid; 56% of reads
Chemistry: Dye-terminator Big Dye; 44% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 174068 bases at least Q40
Consensus quality: 174122 bases at least Q30
Consensus quality: 174133 bases at least Q20
Insert size: 174133; sum-of-contigs
Quality coverage: 7.7x in Q20 bases; sum-of-contigs
-----
Overlapping Sequences:
5': RP11-21287 AC090114 32817-bp clone overlap
3': RP11-274821 (UWGC:djs180) AC018638 115239-bp clone overlap
-----
Sequence Quality Assessment:

```

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation: This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

BcoRI				HindIII				BglII			
SeqDerMap	FgPrnt	SeqDerMap	FgPrnt	SeqDerMap	FgPrnt	SeqDerMap	FgPrnt	SeqDerMap	FgPrnt	SeqDerMap	FgPrnt
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
8696	8841	808	799	3746	3589						
-----	-----	-----	-----	-----	-----						
6	<800	6382	6526	2067	2136						
-----	-----	-----	-----	-----	-----						
4135	4154	512	<800	9001	9123						
-----	-----	-----	-----	-----	-----						
9341	9257	449	<800	4328	4355						
-----	-----	-----	-----	-----	-----						
962	970	6297	6216	1041	1054						
-----	-----	-----	-----	-----	-----						
3828	3866	14754	14447	2308	2467						
-----	-----	-----	-----	-----	-----						
8990	8841	116	<800	931	890						
-----	-----	-----	-----	-----	-----						
1284	1261	1855	1981	10550	10542						
-----	-----	-----	-----	-----	-----						
4271	4259	7682	8065	3920	4229						
-----	-----	-----	-----	-----	-----						
4644	4639	5348	5312	630	<800						
-----	-----	-----	-----	-----	-----						
1874	1837	3865	3811	6869	6855						
-----	-----	-----	-----	-----	-----						
5493	5514	3113	3148	404	<800						
-----	-----	-----	-----	-----	-----						
8821	8841	242	<800	1878	1877						
-----	-----	-----	-----	-----	-----						
3234	3211	1512	1518	8398	8427						
-----	-----	-----	-----	-----	-----						
292	<800	7335	7311	1322	1326						
-----	-----	-----	-----	-----	-----						
1725	1714	2808	2839	687	<800						
-----	-----	-----	-----	-----	-----						
1059	1061	579	<800	7972	8003						
-----	-----	-----	-----	-----	-----						
2055	2073	10626	10446	5685	5677						
-----	-----	-----	-----	-----	-----						
8123	8154	980	986	976	890						
-----	-----	-----	-----	-----	-----						
1808	1837	8983	9103	6594	6629						



all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

ECORI NsiI BglII

SeqDerMap	FgrPrnt	SeqDerMap	FgrPrnt	SeqDerMap	FgrPrnt
8696	8266	20007	20170	2589	2562
6	<800	2189	2217	2067	2086
1059	1053	2222	2217	7030	7221
2055	2050	2633	2650	976	977
8123	8266	334	<800	6594	6600
1808	1818	226	<800	3861	3902
1510	1594	2433	2455	7222	7221
4909	4919	9830	9904	2355	2377
4590	4670	1941	1898	11058	10963
1889	1818	16117	16057	883	869
4920	4919	3548	3575	101	<800
18867	18788	11972	12004	869	869
11516	11684	2958	2952	3181	3204
5915	5970	6864	6910	10120	10205
8939	9018	5934	6024	323	<800
92	<800	917	919	38	<800
3796	3863	2121	2109	1864	1875
1363	1320	152	<800	13497	13351
2077	2050	3177	3223	835	869
2559	2583	2646	2650	2150	2086
3524	3517	407	<800	6576	6600
8210	8266	1039	1040	20897	21438
6569	6560	1797	1791	3378	3384
96	<800	3908	3915	546	<800
1617	1689	12688	12552	9754	9698

4908	4670	5248	5223	3553	3552
2070	2050	1815	1791	466	<800
2757	2795	534	<800	1777	1758
2687	2681	1753	1741	1930	1875
6157	6146	3091	3090	687	<800
4425	4464	13176	13152	6313	6264
4631	4670	3691	3697	2278	2285
1477	1453	6125	6024	4274	4304
1717	1818	3694	3697	1202	1189
4810	4670	6063	6024	2628	2654
1089	1053	2204	2217	158	<800
2871	2890	478	<800	2351	2377
13319	13333	17790	17691	4869	4776
12452	12521	1852	1846	3202	3204
1459	1453			2047	2086
4040	4012			21915	21438
2102	2086				
1902	1875				
2994	2992				
162	<800				

FEATURES

source	Location/Qualifiers
1..176875	/organism="Homo sapiens"
/db_xref="taxon:9606"	
/chromosome="7"	
/clone="RP11-274B21"	
/sex="male"	
/tissue type="blood"	
/clone_lib="RPCI human BAC library 11"	
5530..5596	/standard_name="SWGS3556"
misc_feature	9059..9141
	/note="Single subclone region"
STS	12255..12543
	/standard_name="SWGS1325"
STS	complement(17643..17844)
	/standard_name="SWGS3779"
STS	complement(100521..100620)
	/standard_name="SWGS2227"
STS	complement(135945..136020)
	/standard_name="SWGS3427"

BASE COUNT 50911 a 38997 c 39737 g 47230 t

ORIGIN

Query Match 62.4%; Score 25.6; DB 9; Length 176875;  
Best Local Similarity 77.5%; Pred. No. 12;  
Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
QY 2 CAGTTGCTGAGAGCTGACATACAGGACTGCTGAGAC 41  
DB 149241 CAGCTACTGGAGGCTGACATGTTAGGACTGCTTGACC 149280

```

SULT 14
139396
TUS
INITIATION 178451 bp DNA linear PRI 28-FEB-2002
Human DNA sequence from clone RP11-258C19 on chromosome
Xp11.21-11.23, complete sequence.
ACCESSION AL139396
KEYWORDS HTG.
SOURCE Homo sapiens
ORGANISM human.
REFERENCE 1 (bases 1 to 178451)
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Chapman, J.
TITLE Direct Submission
JOURNAL Submitted (28-FEB-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT On Mar 1, 2002 this sequence version replaced gi:10862725.
During difference assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome X, constructed by the Sanger Centre Chromosome X Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/ChrX
This sequence is the entire insert of clone RP11-258C19 This
sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. RP11-258C19 is from
the library RPCI-11.1 constructed by the group of Pieter de Jong.
For further details see
http://www.choxi.org/bacpac/home.htm
VECTOR: pBACe3.6.
FEATURES
Location/Qualifiers
1..178451
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="p11.21-11.23"
/clone="RP11-258C19"
/clone_lib="RPCI-11.1"
178349..178451
/notes="Single clone region. Assembly confirmed by
restriction digest data"
SE COUNT 48101 a 42307 c 40594 g 47449 t
IGIN
Query Match 62.4%; Score 25.6; DB 9; Length 178451;
Best Local Similarity 77.5%; Pred. No. 12;
Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
2 CAGTGTCTGAGAGCTGTGATACAGGAGGAGCTGCTGAGAC 41
|||||
61898 CAGCTGCTTAGGAGCTGTGAGGAGGAGGAGGAGCTGCTGAGCC 61937
|||||
SULT 15
011755/c
TUS
AC011755 179110 bp DNA linear PRI 07-NOV-2001
Homo sapiens BAC clone RP11-575M22 from 2, complete sequence.
AC011755
AC011755.7 GI:14140349
HTG.
Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 179110)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Sulston, J.E. and Waterston, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 179110)
Martinka, S., Abbott, A. and Boyer, E.
TITLE The sequence of Homo sapiens BAC clone RP11-575M22
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 179110)
Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (13-OCT-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 179110)
Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (17-MAY-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 5 (bases 1 to 179110)
Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (07-NOV-2001) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On May 17, 2001 this sequence version replaced gi:13270817.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
Center project name: H_NH0575M22
-----
NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.
This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.
MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.
McPherson, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see http://genome.wustl.edu/gsc
SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male
donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E.,
Tateno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
and coworkers at the Roswell Park Cancer Institute
(http://bacpac.med.buffalo.edu)

```

VECTOR: pBACe3.6  
 NEIGHBORING SEQUENCE INFORMATION:  
 The clone sequenced to the left is RP11-9313, 200 bp overlap; the  
 clone sequenced to the right is RP11-449L24, 200 bp overlap.  
 Actual start of this clone is at base position 195 of RP11-575M22;  
 actual end is at base position 178916 of RP11-575M22.

RP11-575M22 contains single stranded regions from 78912 to 79086  
 and 123395 to 123463.

FEATURES  
 source  
 1. 179110  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="2"  
 /map="2"  
 /clone="RP11-575M22"  
 /clone\_lib="RPCL-11"  
 105..127  
 /rpt\_family="AT-rich"  
 258..338  
 /rpt\_family="MER2\_type"  
 467..866  
 /rpt\_family="MaLR"  
 1757..1831  
 /rpt\_family="MIR"  
 1917..2043  
 /rpt\_family="MIR"  
 3551..3856  
 /rpt\_family="L1"  
 3918..4005  
 /rpt\_family="L2"  
 4062..4136  
 /rpt\_family="L1"  
 4390..4571  
 /rpt\_family="MER1\_type"  
 5831..6116  
 /rpt\_family="L1"  
 7659..7725  
 /rpt\_family="(TTCA)n"  
 7743..8144  
 /rpt\_family="L2"  
 8337..8482  
 /rpt\_family="MIR"  
 8869..9083  
 /rpt\_family="MER1\_type"  
 9971..10095  
 /rpt\_family="MIR"  
 10111..10224  
 /rpt\_family="L2"  
 10309..10605  
 /rpt\_family="Alu"  
 11779..11927  
 /rpt\_family="L2"  
 12527..12779  
 /rpt\_family="MER1\_type"  
 13379..13629  
 /rpt\_family="MER1\_type"  
 14028..14161  
 /rpt\_family="Alu"  
 14192..14619  
 /rpt\_family="MaLR"  
 14827..15204  
 /rpt\_family="MaLR"  
 15380..15499  
 /rpt\_family="L2"  
 15548..15849  
 /rpt\_family="Alu"  
 16029..16444  
 /rpt\_family="MaLR"  
 16445..16630  
 /rpt\_family="MER2\_type"  
 16957..17105  
 /rpt\_family="MIR"

repeat\_region 18513..19133  
 /rpt\_family="MaLR"  
 repeat\_region 19308..19523  
 /rpt\_family="MaLR"  
 repeat\_region 19557..19649  
 /rpt\_family="MaLR"  
 repeat\_region 19767..19831  
 /rpt\_family="Mariner"  
 repeat\_region 20151..20482  
 /rpt\_family="Alu"  
 21962..22347  
 /note="similar to EST AW751954 (NID:G7666990)"  
 22130..22493  
 /note="similar to EST T07269 (NID:G318418)"  
 23745..23932  
 /rpt\_family="MER1\_type"  
 24295..24451  
 /rpt\_family="MIR"  
 24818..25577  
 /rpt\_family="L1"  
 26071..26387  
 /rpt\_family="Alu"  
 26687..26732  
 /rpt\_family="(CA)n"  
 26806..26954  
 /rpt\_family="MIR"  
 27219..27363  
 /rpt\_family="MIR"  
 27454..27722  
 /rpt\_family="Alu"  
 29988..30259  
 /rpt\_family="Alu"  
 30749..30844  
 /rpt\_family="MIR"  
 31573..31690  
 /rpt\_family="L2"  
 31893..32044  
 /rpt\_family="CR1"  
 32410..32433  
 /rpt\_family="(TGAA)n"  
 32544..32924  
 /rpt\_family="L1"  
 33116..33387  
 /note="similar to EST BE082071 (NID:G8472367)"  
 33494..33516  
 /rpt\_family="AT-rich"  
 34711..35101  
 /rpt\_family="MaLR"  
 36087..36139  
 /rpt\_family="MIR"  
 36140..36428  
 /rpt\_family="Alu"  
 36429..36472

Query Match 62.4%; Score 25.6; DB 9; Length 179110;  
 Best Local Similarity 77.5%; Pred. No. 12;  
 Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 2 CAGTTGCTGAGAGCCTGACATACCGAGGACTGCGCTGAGAC 41  
 |||||  
 Db 149688 CAGTTACTCAGGAGGCTGACATAAAGGATTGCTTGAGCC 149649

Search completed: July 22, 2003, 13:43:32  
 Job time : 992 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

nucleic - nucleic search, using sw model

, on: July 22, 2003, 11:16:53 ; Search time 170.5 Seconds  
(without alignments)  
541.536 Million cell updates/sec

le: U95626-C-AT-42723

ect score: 41

uence: 1 tcagttgtgagactgta.....ataccagactgctgagac 41

ring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

rchd: 2185239 seqs, 112599159 residues

al number of hits satisfying chosen parameters: 4370478

imum DB seq length: 0

ximum DB seq length: 2000000000

st-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

labase : N Genseq 101002.\*

- 1: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1980.DAT.\*
- 2: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1981.DAT.\*
- 3: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1982.DAT.\*
- 4: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1983.DAT.\*
- 5: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1984.DAT.\*
- 6: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1985.DAT.\*
- 7: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1986.DAT.\*
- 8: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1987.DAT.\*
- 9: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1988.DAT.\*
- 10: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1989.DAT.\*
- 11: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1990.DAT.\*
- 12: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1991.DAT.\*
- 13: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1992.DAT.\*
- 14: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1993.DAT.\*
- 15: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1994.DAT.\*
- 16: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1995.DAT.\*
- 17: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1996.DAT.\*
- 18: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1997.DAT.\*
- 19: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1998.DAT.\*
- 20: /SIDS2/gcgdata/genseq/genseqn-emb1/NA1999.DAT.\*
- 21: /SIDS2/gcgdata/genseq/genseqn-emb1/NA2000.DAT.\*
- 22: /SIDS2/gcgdata/genseq/genseqn-emb1/NA2001A.DAT.\*
- 23: /SIDS2/gcgdata/genseq/genseqn-emb1/NA2001B.DAT.\*
- 24: /SIDS2/gcgdata/genseq/genseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Query	Score	Match	Length	ID	Description
1	41	100.0	143068	21	AAF21105
2	41	100.0	143068	21	AAF21272
3	41	100.0	143068	21	AAA34983
4	41	100.0	143068	21	AAA35150
5	41	100.0	143068	21	AAA35150
6	41	100.0	143068	24	ABL68124
7	41	100.0	149412	21	AAA35151
8	25.8	62.9	633	22	AAF21273
9	25.6	62.4	1001	22	AAF181867

10	25.2	61.5	604	22	AAI20925
11	25.2	61.5	604	22	AAI23598
C 12	24	58.5	349	22	ABAI4374
C 13	24	58.5	397	22	AAI91318
C 14	24	58.5	935	24	AAI16521
C 15	24	58.5	992	23	AAI75085
C 16	24	58.5	1624	22	ABAI5771
C 17	24	58.5	1624	22	AAK73388
C 18	24	58.5	3565	16	AAO84051
C 19	24	58.5	4749	22	AAH18519
C 20	24	58.5	10528	24	ABL32334
C 21	24	58.5	17335	23	ABK42393
C 22	24	58.5	19882	23	ABK42394
C 23	23.4	57.1	413	22	AAI81647
C 24	23.4	57.1	10528	24	ABL32335
C 25	23.4	57.1	38374	24	ABN96966
C 26	23.4	57.1	38374	24	ABL68363
C 27	23.4	57.1	38374	24	ABL68364
C 28	23.4	57.1	38374	24	ABL68824
C 29	23	56.1	178896	24	ABQ88146
C 30	22.8	55.6	1912	20	AAV83767
C 31	22.6	55.1	908	23	AAI84558
C 32	22.6	55.1	75384	22	AAK85590
C 33	22.4	54.6	329	22	AAI91783
C 34	22.4	54.6	360	22	AAI82165
C 35	22.4	54.6	387	22	AAK67876
C 36	22.4	54.6	395	22	AAI82659
C 37	22.4	54.6	441	23	ABV56385
C 38	22.4	54.6	445	22	AAK86122
C 39	22.4	54.6	445	22	AAK86123
C 40	22.4	54.6	445	22	AAK86124
C 41	22.4	54.6	450	22	AAI32604
C 42	22.4	54.6	564	23	AAI71886
C 43	22.4	54.6	574	23	ABV30511
C 44	22.4	54.6	574	23	ABV39484
C 45	22.4	54.6	799	23	ABV09335

#### ALIGNMENTS

RESULT 1

AAF21105

ID AAF21105 standard; DNA; 143068 BP.

XX AAF21105;

AC AAF21105;

DT 14-MAR-2001 (first entry)

XX Human low adenosine antisense oligonucleotide related sequence #2672.

DE Low adenosine antisense oligonucleotide; phosphorothioate; allergy;  
KW human; airway disorder; bronchoconstriction; lung inflammation;  
KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;  
KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;  
KW respiratory obstruction; pulmonary obstruction; impeded respiration;  
KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;  
KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;  
KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;  
KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;  
KW cancer; ss.

OS Homo sapiens.

XX WO200062736-A2.

FN 26-OCT-2000.

PD 24-MAR-2000; 2000WO-US08020.

PF 06-APR-1999; 99US-0127958.

XX (UYEC-) UNIV EAST CAROLINA.

PA



(NYCE/) NYCE J W.

Nyce JW;

WFI; 2000-679539/66.

Low adenosine (A) content antisense oligonucleotides which do not trigger adenosine receptors during metabolism, useful e.g. for treating cancers and respiratory obstructions -

Disclosure; Page 924-957; 1592pp; English.

The present invention describes low adenosine (A) content antisense oligonucleotides and compositions (I) comprising them. In the antisense oligonucleotides the A is replaced by a 'Universal' or alternative base. (I) can have respiratory, bronchodilator, antiinflammatory, analgesic, immunosuppressive, antiasthmatic, hypotensive and cytostatic activities. The antisense oligonucleotides and (I) can be used to down-regulate the expression and/or activity of target polypeptides associated with lung/respiratory disorders and malignancies, such as stimulating and activating peptide factors and transmitters, transcription factors, immunoglobulins and antibodies, antibody receptors, cytokines and chemokines, endogenously produced specific and non-specific enzymes, binding proteins, adhesion molecules and their receptors, cytokine and chemokine receptors, adenosine receptors, bradykinin receptors, central nervous system (CNS) and peripheral nervous and non-nervous system receptors, CNS and peripheral nervous and non-nervous system transmitters, defensins, growth factors, vasoactive peptides and receptors, binding proteins and malignancy associated proteins. The antisense oligonucleotides may be used in this way to treat disorders including respiratory obstruction (especially pulmonary obstruction and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or surfactant hypoproduction which are associated with a disease or condition selected from pulmonary vasoconstriction, inflammation, allergies, asthma, impeded respiration, respiratory distress syndrome (ARDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), pulmonary transplantation rejection, pulmonary infections, bronchitis, and/or cancer. AAF18434 to AAF21543 represent human polynucleotide fragments and antisense oligonucleotides used in the exemplification of the present invention.

Sequence 143068 BP; 41194 A; 30122 C; 32403 G; 39349 T; 0 other;

Query Match 100.0%; Score 41; DB 21; Length 143068;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TCAGTTCGTGAGAGCCTGCATACACAGGACTGCCTGAGAC 41

42703 TCAGTTCGTGAGAGCCTGCATACACAGGACTGCCTGAGAC 42743

SU1T 2

F21272

AAF21272 standard; DNA; 143068 BP.

AAF21272;

14-MAR-2001 (first entry)

Human low adenosine antisense oligonucleotide related sequence #2839.

Low adenosine antisense oligonucleotide; phosphorothioate; allergy; human; airway disorder; bronchoconstriction; lung inflammation; surfactant depletion; respiratory bronchodilator; antiinflammatory; immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic; respiratory obstruction; pulmonary obstruction; impeded respiration; surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS; respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis; pulmonary hypertension; emphysema; pulmonary transplantation rejection; chronic obstructive pulmonary disease; pulmonary infection; bronchitis; cancer; ss.

XX OS Homo sapiens.  
XX PN WO200062736-A2.  
XX PD 26-OCT-2000.  
XX PF 24-MAR-2000; 2000WO-US08020.  
XX PR 06-APR-1999; 99US-0127958.  
XX PA (UYEC-) UNIV EAST CAROLINA.  
XX PI (NYCE/) NYCE J W.  
XX NYce JW;  
XX WFI; 2000-679539/66.  
XX Low adenosine (A) content antisense oligonucleotides which do not trigger adenosine receptors during metabolism, useful e.g. for treating cancers and respiratory obstructions -  
XX Disclosure; Page 1186-1219; 1592pp; English.

The present invention describes low adenosine (A) content antisense oligonucleotides and compositions (I) comprising them. In the antisense oligonucleotides the A is replaced by a 'Universal' or alternative base. (I) can have respiratory, bronchodilator, antiinflammatory, analgesic, immunosuppressive, antiasthmatic, hypotensive and cytostatic activities. The antisense oligonucleotides and (I) can be used to down-regulate the expression and/or activity of target polypeptides associated with lung/respiratory disorders and malignancies, such as stimulating and activating peptide factors and transmitters, transcription factors, immunoglobulins and antibodies, antibody receptors, cytokines and chemokines, endogenously produced specific and non-specific enzymes, binding proteins, adhesion molecules and their receptors, cytokine and chemokine receptors, adenosine receptors, bradykinin receptors, central nervous system (CNS) and peripheral nervous and non-nervous system receptors, CNS and peripheral nervous and non-nervous system transmitters, defensins, growth factors, vasoactive peptides and receptors, binding proteins and malignancy associated proteins. The antisense oligonucleotides may be used in this way to treat disorders including respiratory obstruction (especially pulmonary obstruction and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or surfactant hypoproduction which are associated with a disease or condition selected from pulmonary vasoconstriction, inflammation, allergies, asthma, impeded respiration, respiratory distress syndrome (ARDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), pulmonary transplantation rejection, pulmonary infections, bronchitis, and/or cancer. AAF18434 to AAF21543 represent human polynucleotide fragments and antisense oligonucleotides used in the exemplification of the present invention.

Sequence 143068 BP; 41194 A; 30122 C; 32403 G; 39349 T; 0 other;

Query Match 100.0%; Score 41; DB 21; Length 143068;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCAGTTCGTGAGAGCCTGCATACACAGGACTGCCTGAGAC 41

Db 42703 TCAGTTCGTGAGAGCCTGCATACACAGGACTGCCTGAGAC 42743

RESULT 3

AAA34983

ID AAA34983 standard; DNA; 143068 BP.

XX AC AAA34983;

XX DT 28-JUL-2000 (first entry)

XX

Human adenosine receptor related polynucleotide SEQ ID NO:2672.

Human; adenosine receptor; low adenosine antisense oligonucleotide; phosphorothioate; impaired respiration; inflammation; allergy; allergic disease; bronchoconstriction; inhibitor; anti-inflammatory; antiallergic; antisthmatic; cytosstatic; analgesic; impaired airway; lung disease; ischaemic condition; pulmonary vasoconstriction; asthma; respiratory distress syndrome; pain; cystic fibrosis; emphysema; pulmonary hypertension; chronic obstructive pulmonary disease; COPD; cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.

Homo sapiens.

W0200009525-A2.

24-FEB-2000.

03-AUG-1999; 99WO-US17712.

03-AUG-1998; 98US-0095212.

(UYEC-) UNIV EAST CAROLINA.

Nyce JW;

WPI; 2000-205971/18.

New antisense oligonucleotides useful for treating e.g. pulmonary vasoconstriction, inflammation, allergies, asthma, hypertension, bronchitis, emphysema, respiratory distress syndrome, ischemia or cancers -

Disclosure; Page 851-882; 1343pp; English.

The present invention describes a new composition comprising an antisense oligonucleotide (ON) with low adenosine (up to 15%), which targets nucleic acids involved in bronchoconstriction, allergies, and/or inflammation. The ON can have anti-inflammatory, antiallergic, antisthmatic, cytosstatic and analgesic activities. The compositions are useful for the treatment of diseases associated with inflammation,

impaired airways, including lung disease and diseases whose secondary effects afflict the lungs of a subject. They can be used for treating e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma, impaired respiration, respiratory distress syndrome, pain, cystic fibrosis, pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), and cancers such as leukaemias, lymphomas, carcinomas, and cancers which may metastasise to the lungs, including breast and prostate cancer. The reduction of the adenosine content of the ONs reduces side effects. The A-containing ONs break down with the release of deoxyadenosine which activates adenosine receptors causing bronchoconstriction and inflammation. AAA32313 to AAA35112 represent the nucleotide sequences given in the sequence listing from the present invention, which correspond to SEQ ID NO:1 to 2815, and then the last 185 sequences are also called SEQ ID NO:1 to 185, but the sequences differ from the previously named sequences. SEQ ID NO:11 to 1680 (AAA32323 to AAA33992) are specifically claimed ONs from the present invention. N.B. Sequences given in the disclosure of the present invention do not match up with their corresponding SEQ ID NO: sequences given in the sequence listing.

Sequence 143068 BP; 41194 A; 30122 C; 32402 G; 39350 T; 0 other;

Query Match 100.0%; Score 41; DB 21; Length 143068;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TCAGTTGCTGAGAGCCTGCATACACAGGACTGCTGAGAC 41

|||||

42703 TCAGTTGCTGAGAGCCTGCATACACAGGACTGCTGAGAC 42743

ULT 4  
 15150

AA35150 standard; DNA; 143068 BP.  
 AC AA35150;  
 DT 28-JUL-2000 (first entry)  
 DE Human adenosine receptor related polynucleotide 2nd SEQ ID NO:24.  
 KW Human; adenosine receptor; low adenosine antisense oligonucleotide; phosphorothioate; impaired respiration; inflammation; allergy; allergic disease; bronchoconstriction; inhibitor; anti-inflammatory; antiallergic; antisthmatic; cytosstatic; analgesic; impaired airway; lung disease; ischaemic condition; pulmonary vasoconstriction; asthma; respiratory distress syndrome; pain; cystic fibrosis; emphysema; pulmonary hypertension; chronic obstructive pulmonary disease; COPD; cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.  
 KW Homo sapiens.  
 OS W0200009525-A2.  
 XX 24-FEB-2000.  
 XX 03-AUG-1999; 99WO-US17712.  
 PF 03-AUG-1998; 98US-0095212.  
 PR (UYEC-) UNIV EAST CAROLINA.  
 PA Nyce JW;  
 PI WPI; 2000-205971/18.  
 DR XX  
 PT New antisense oligonucleotides useful for treating e.g. pulmonary vasoconstriction, inflammation, allergies, asthma, hypertension, bronchitis, emphysema, respiratory distress syndrome, ischemia or cancers -  
 PT Disclosure; Page 1106-1138; 1343pp; English.  
 XX The present invention describes a new composition comprising an antisense oligonucleotide (ON) with low adenosine (up to 15%), which targets nucleic acids involved in bronchoconstriction, allergies, and/or inflammation. The ON can have anti-inflammatory, antiallergic, antisthmatic, cytosstatic and analgesic activities. The compositions are useful for the treatment of diseases associated with inflammation, impaired airways, including lung disease and diseases whose secondary effects afflict the lungs of a subject. They can be used for treating e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma, impaired respiration, respiratory distress syndrome, pain, cystic fibrosis, pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), and cancers such as leukaemias, lymphomas, carcinomas, and cancers which may metastasise to the lungs, including breast and prostate cancer. The reduction of the adenosine content of the ONs reduces side effects. The A-containing ONs break down with the release of deoxyadenosine which activates adenosine receptors causing bronchoconstriction and inflammation. AAA32313 to AAA35112 represent the nucleotide sequences given in the sequence listing from the present invention, which correspond to SEQ ID NO:1 to 2815, and then the last 185 sequences are also called SEQ ID NO:1 to 185, but the sequences differ from the previously named sequences. SEQ ID NO:11 to 1680 (AAA32323 to AAA33992) are specifically claimed ONs from the present invention. N.B. Sequences given in the disclosure of the present invention do not match up with their corresponding SEQ ID NO: sequences given in the sequence listing.

Sequence 143068 BP; 41194 A; 30126 C; 32402 G; 39346 T; 0 other;  
 Query Match 100.0%; Score 41; DB 21; Length 143068;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TCAGTTGCTGAGAGCCTGCATACACAGGACTGCTGAGAC 41

|||||TCAGTTGCTGAGAGCCTGACATACACGAGACTGCTGAGAC 42743

|||||  
JUL 5  
68124  
ABU68124 standard; DNA; 143068 BP.

ABU68124;

15-MAY-2002 (first entry)

Ovary cancer related gene sequence SEQ ID NO:6461.

Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
gene; ds.

Homo sapiens.

WO200194629-A2.

13-DEC-2001.

30-MAY-2001; 2001WO-US10838.

05-JUN-2000; 2000US-209473P.

05-JUN-2000; 2000US-209531P.

18-SEP-2000; 2000US-233133P.

18-SEP-2000; 2000US-233617P.

20-SEP-2000; 2000US-234009P.

20-SEP-2000; 2000US-234034P.

20-SEP-2000; 2000US-234052P.

22-SEP-2000; 2000US-234509P.

22-SEP-2000; 2000US-234567P.

25-SEP-2000; 2000US-234923P.

25-SEP-2000; 2000US-234924P.

25-SEP-2000; 2000US-235077P.

25-SEP-2000; 2000US-235082P.

25-SEP-2000; 2000US-235134P.

25-SEP-2000; 2000US-235280P.

26-SEP-2000; 2000US-235637P.

26-SEP-2000; 2000US-235638P.

27-SEP-2000; 2000US-235711P.

27-SEP-2000; 2000US-235720P.

27-SEP-2000; 2000US-235840P.

27-SEP-2000; 2000US-235863P.

28-SEP-2000; 2000US-236028P.

28-SEP-2000; 2000US-236032P.

28-SEP-2000; 2000US-236033P.

28-SEP-2000; 2000US-236034P.

28-SEP-2000; 2000US-236109P.

PI Soppet DR, Weaver Z;  
XX WPI; 2002-188264/24.  
XX  
PT Screening for anti-neoplastic agent involves exposing cells to a  
PT chemical agent to be tested for anti-neoplastic activity, and  
PT determining a change in expression of a gene of a signature gene set -  
XX  
XX Claim 1; SEQ ID 6461; 44pp; English.  
XX

CC The present invention describes a method (M1) for screening for an  
CC anti-neoplastic agent. The method involves exposing cells to a chemical  
CC agent to be tested for anti-neoplastic activity, determining a change in  
CC expression of at least one gene (I) of a signature gene set, where (I)  
CC comprises a sequence (S) selected from 8447 sequences (given in ABU61664  
CC to ABU70110), or is at least 95% identical to (S), where a change in  
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
CC activity and can be used in gene therapy. M1 can be used for screening  
CC an anti-neoplastic agent, and can be used for producing a product which  
CC is the data collected with respect to the anti-neoplastic agent as a  
CC result of M1, and the data is sufficient to convey the chemical  
CC structure and/or properties of the agent. M1 can be used in the  
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,  
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
CC carcinoma, papillary carcinoma and Wilm's tumour.  
XX

SQ Sequence 143068 BP; 41194 A; 30122 C; 32403 G; 39349 T; 0 other;  
Query Match 100.0%; Score 41; DB 24; Length 143068;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TCAGTTGCTGAGAGCCTGACATACACGAGACTGCTGAGAC 41  
DB 42703 TCAGTTGCTGAGAGCCTGACATACACGAGACTGCTGAGAC 42743

RESULT 6

AAA35151

ID AAA35151 standard; DNA; 149412 BP.

XX AAA35151;

XX 28-JUN-2000 (first entry)

DE Human adenosine receptor related polynucleotide 2nd SEQ ID NO:25.

XX Human; adenosine receptor; low adenosine antisense oligonucleotide;  
KW phosphorothioate; impaired respiration; inflammation; allergy;  
KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;  
KW antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;  
KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;  
KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;  
KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;  
KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.

OS Homo sapiens.

XX WO200009525-A2.

XX 24-FEB-2000.

PD 03-AUG-1999; 99WO-US17712.

XX 03-AUG-1998; 98US-0095212.

XX (UYEC-) UNIV EAST CAROLINA.

XX Nyce JW;

XX WPI; 2000-205971/18.

Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;

(AVAL-) AVALON PHARM.

New antisense oligonucleotides useful for treating e.g. pulmonary vasoconstriction, inflammation, allergies, asthma, hypertension, bronchitis, emphysema, respiratory distress syndrome, ischemia or cancers -

Disclosure; Page 1138-1171; 1343pp; English.

The present invention describes a new composition comprising an antisense oligonucleotide (ON) with low adenosine (up to 15%), which targets nucleic acids involved in bronchoconstriction, allergies, and/or inflammation. The ON can have antiinflammatory, antiallergic, antiasthmatic, cytostatic and analgesic activities. The compositions are useful for the treatment of diseases associated with inflammation, impaired airways, including lung disease and diseases whose secondary effects afflict the lungs of a subject. They can be used for treating e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma, impaired respiration, respiratory distress syndrome, pain, cystic fibrosis, pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), and cancers such as leukaemias, lymphomas, carcinomas, and cancers which may metastasise to the lungs, including breast and prostate cancer. The reduction of the adenosine content of the ONs reduces side effects. The A-containing ONs break down with the release of deoxyadenosine which activates adenosine receptors causing bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the nucleotide sequences given in the sequence listing from the present invention, which correspond to SEQ ID NO:1 to 2815, and then the last 185 sequences are also called SEQ ID NO:1 to 185, but the sequences differ from the previously named sequences. SEQ ID NO:11 to 1680 (AAA32323 to AAA33992) are specifically claimed ONs from the present invention. N.B. Sequences given in the disclosure of the present invention do not match up with their corresponding SEQ ID NO: sequences given in the sequence listing.

Sequence 149412 BP; 43049 A; 31388 C; 33852 G; 41123 T; 0 other;

Query Match 100.0%; Score 41; DB 21; Length 149412;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TCAGTTGCTGAGAGCCTGACATACCGAGCTGCTGAGAC 41  
|||||  
49047 TCAGTTGCTGAGAGCCTGACATACCGAGCTGCTGAGAC 49087

FULT 7

21273  
AAF21273 standard; DNA; 152740 BP.

AAF21273;

14-MAR-2001 (first entry)

Human low adenosine antisense oligonucleotide related sequence #2840.

Low adenosine antisense oligonucleotide; phosphorothioate; allergy; human; airway disorder; bronchoconstriction; lung inflammation; surfactant depletion; respiratory; bronchodilator; antiinflammatory; immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic; respiratory obstruction; pulmonary vasoconstriction; impaired respiration; surfactant hypoproduction; pulmonary obstruction; asthma; RDS; respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis; pulmonary hypertension; emphysema; pulmonary transplantation rejection; chronic obstructive pulmonary disease; pulmonary infection; bronchitis; cancer; ss.

Homo sapiens.

WO200062736-A2.

26-OCT-2000.

24-MAR-2000; 2000WO-US08020.

XX PR 06-APR-1999; 99US-0127958.  
XX PA (UYEC-) UNIV EAST CAROLINA.  
XX PA (NYCE/) NYCE J W.  
XX PI Nyce JW;  
XX WI; 2000-679539/66.  
DR WPI; 2000-679539/66.

XX Low adenosine (A) content antisense oligonucleotides which do not  
PT trigger adenosine receptors during metabolism, useful e.g. for treating  
PT cancers and respiratory obstructions -

XX Disclosure; Page 1219-1254; 1592pp; English.

XX The present invention describes low adenosine (A) content antisense  
CC oligonucleotides and compositions (I) comprising them. In the antisense  
CC oligonucleotides the A is replaced by a 'Universal' or alternative base.  
CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,  
CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.  
CC The antisense oligonucleotides and (I) can be used to down-regulate the  
CC expression and/or activity of target polypeptides associated with  
CC lung/respiratory disorders and malignancies, such as stimulating and  
CC activating peptide factors and transmitters, transcription factors,  
CC immunoglobulins and antibodies, antibody receptors, cytokines and  
CC chemokines, endogenously produced specific and non-specific enzymes,  
CC binding proteins, adhesion molecules and their receptors, cytokine and  
CC chemokine receptors, adenosine receptors, bradykinin receptors, central  
CC nervous system (CNS) and peripheral nervous and non-nervous system  
CC receptors, CNS and peripheral nervous and non-nervous system peptide  
CC transmitters, defensins, growth factors, vasoactive peptides and  
CC receptors, binding proteins and malignancy associated proteins. The  
CC antisense oligonucleotides may be used in this way to treat disorders  
CC including respiratory obstruction (especially pulmonary obstruction  
CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies)  
CC and/or surfactant hypoproduction which are associated with a disease or  
CC condition selected from pulmonary vasoconstriction, inflammation,  
CC allergies, asthma, impaired respiration, respiratory distress syndrome  
CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary  
CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),  
CC pulmonary transplantation rejection, pulmonary infections, bronchitis,  
CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide  
CC fragments and antisense oligonucleotides used in the exemplification of  
CC the present invention.

XX SQ Sequence 152740 BP; 44169 A; 32023 C; 34549 G; 41999 T; 0 other;

Query Match 100.0%; Score 41; DB 21; Length 152740;  
Best Local Similarity 100.0%; Pred. No. 1.6e-06;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAGTTGCTGAGAGCCTGACATACCGAGCTGCTGAGAC 41  
|||||  
Db 49047 TCAGTTGCTGAGAGCCTGACATACCGAGCTGCTGAGAC 49087

RESULT 8

AA114736

XX AAL14736 standard; cDNA; 633 BP.

XX AAL14736;

DT 07-DEC-2001 (first entry)

XX Human breast cancer expressed polynucleotide 7193.

XX Human; breast cancer; cell marker; cytostatic; ss.

XX Homo sapiens.

XX WO200151628-A2.

19-JUL-2001.  
 10-JAN-2001; 2001WO-US00798.  
 14-JAN-2000; 2000US-0176077.  
 14-MAR-2000; 2000US-0189167.  
 24-MAR-2000; 2000US-0192099.  
 29-MAR-2000; 2000US-0193480.  
 15-MAY-2000; 2000US-0205230.  
 09-JUN-2000; 2000US-0211315.  
 25-JUL-2000; 2000US-0220534.  
 (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 Lillie J, Xu Y, Wang Y, Steinmann K;  
 WPI; 2001-451856/48.  
 New peptide useful as a marker for the diagnosis of breast cancer -  
 Claim 1; Page 1300; 3695pp; English.  
 The invention relates to human breast cancer expressed polynucleotides (AAI07544-AAI26789) and methods of assessing whether a patient is afflicted with breast cancer by examining the correlation between the expression of certain markers and the cancerous state of breast cells. The polynucleotides and encoded polypeptides are potential markers for detecting, diagnosing, monitoring, characterising, treating and potentially preventing breast cancer. The polynucleotides and encoded polypeptides are also useful for isolating compounds with cytostatic activity.  
 Sequence 633 BP; 223 A; 92 C; 178 G; 136 T; 4 other;  
 Query Match 62.9%; Score 25.8; DB 22; Length 633;  
 Best Local Similarity 78.9%; Pred. No. 0.75; Indels 0; Gaps 0;  
 Matches 30; Conservative 0; Mismatches 8;  
 1 TCAGTTCCTGAGAGCTGCATACGAGCTGCTGA 38  
 522 TCAGTTCCTGAGAGCTGCATGCTGANGAGATTGCTGA 559  
 SULT 9  
 T81867/c  
 AAI81867 standard; cDNA; 1001 BP.  
 AAI81867;  
 06-NOV-2001 (first entry)  
 Human polynucleotide SEQ ID NO 1927.  
 Human; cytokine; cell proliferation; cell differentiation; gene therapy; vaccine; peptide therapy; stem cell growth factor; haematopoiesis; tissue growth factor; immunomodulatory; cancer; leukaemia; nervous system disorders; arthritis; inflammation; ss.  
 Homo sapiens.  
 WO200164835-A2.  
 07-SEP-2001.  
 26-FEB-2001; 2001WO-US04927.  
 28-FEB-2000; 2000US-0515126.  
 18-MAY-2000; 2000US-0577409.  
 (HYSE-) HYSEQ INC.  
 Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-514838/56.  
 DR P-PSDB; AAO1936.  
 XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing and treating e.g. leukaemia, inflammation and immune disorders -  
 PT  
 PT  
 XX  
 PS Claim 1; SEQ ID NO 1927; 1399pp + Sequence Listing; English.  
 XX  
 CC The invention relates to human polynucleotides (AAI79941-AAI93841) and the encoded proteins (AAO0010-AAO13910) that exhibit activity elating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation.  
 CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 1001 BP; 258 A; 272 C; 239 G; 229 T; 3 other;  
 Query Match 62.4%; Score 25.6; DB 22; Length 1001;  
 Best Local Similarity 77.5%; Pred. No. 1;  
 Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
 QY 2 CAGTTCCTGAGAGCTGCATACGAGCTGCTGAGAC 41  
 DB 908 CAGCTACTGGAAGCTGCATGCTGAGCTGCTGAGCC 869  
 RESULT 10  
 AAL20925  
 ID AAL20925 standard; cDNA; 604 BP.  
 XX  
 AC AAL20925;  
 XX  
 DT 07-DEC-2001 (first entry)  
 XX  
 DE Human breast cancer expressed polynucleotide 13382.  
 XX  
 KW Human; breast cancer; cell marker; cytostatic; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200151628-A2.  
 XX  
 PD 19-JUL-2001.  
 XX  
 PF 10-JAN-2001; 2001WO-US00798.  
 XX  
 PR 14-JAN-2000; 2000US-0176077.  
 PR 14-MAR-2000; 2000US-0189167.  
 PR 24-MAR-2000; 2000US-0192099.  
 PR 29-MAR-2000; 2000US-0193480.  
 PR 15-MAY-2000; 2000US-0205230.  
 PR 09-JUN-2000; 2000US-0211315.  
 PR 25-JUL-2000; 2000US-0220534.  
 XX  
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 XX  
 PI Lillie J, Xu Y, Wang Y, Steinmann K;  
 XX  
 DR WPI; 2001-451856/48.  
 XX  
 PT New peptide useful as a marker for the diagnosis of breast cancer -  
 XX  
 PS Claim 1; Page 2374; 3695pp; English.  
 XX

Very Match 61.5%; Score 25.2; DB 22; Length 604;  
 1st Local Similarity 78.9%; Pred. No. 1.3;

Matches	30;	Conservative	0;	Mismatches	8;	Indels	0;	Gaps	0;
QY	1	TCAGTTGCTGAGAGCCTGACATACCAGGAGTGCTGGA	38						
Db	495	TCAGTGTCTCAGAGCCTGACGTAAGAGAGATTGCTTGA	532						
RESULT 12									
ABAL4274/c									
ID	ABA14274 standard; cDNA; 349 BP.								
XX	AC								
XX	ABAL4274;								
DT	23-JAN-2002 (first entry)								
XX	XX	Human nervous system related polynucleotide SEQ ID NO 3281.							
DE	XX								
XX	XX	Human; neurotropic; neuroprotective; cytostatic; dermatological; virucide;							
KW	KW	immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;							
KW	KW	antiparkinsonian; antischling; antinaemic; antiarthritic; cancer;							
KW	KW	antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;							
KW	KW	antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;							
KW	KW	antiparasitic; cardiac; immune disorder; cardiovascular disorder;							
KW	KW	neurological disease; infection; immunogenic; nephrotropic; vaccine; ss.							
XX	XX	Homo sapiens.							
OS	XX	Homo sapiens.							
PN	XX	WO200159063-A2.							
XX	XX	16-AUG-2001.							
PD	XX								
XX	XX	17-JAN-2001; 2001WO-US01334.							
PF	XX								
PP	XX	31-JAN-2000; 2000US-0179065.							
PR	PR	04-FEB-2000; 2000US-0180628.							
PR	PR	24-FEB-2000; 2000US-0184664.							
PR	PR	02-MAR-2000; 2000US-0186350.							
PR	PR	16-MAR-2000; 2000US-0189874.							
PR	PR	17-MAR-2000; 2000US-0190076.							
PR	PR	18-APR-2000; 2000US-0198123.							
PR	PR	19-MAY-2000; 2000US-0205515.							
PR	PR	07-JUN-2000; 2000US-0209467.							
PR	PR	28-JUN-2000; 2000US-0214886.							
PR	PR	30-JUN-2000; 2000US-0215135.							
PR	PR	07-JUL-2000; 2000US-0216647.							
PR	PR	07-JUL-2000; 2000US-0216880.							
PR	PR	11-JUL-2000; 2000US-0217487.							
PR	PR	11-JUL-2000; 2000US-0217496.							
PR	PR	14-JUL-2000; 2000US-0218290.							
PR	PR	26-JUL-2000; 2000US-0220963.							
PR	PR	26-JUL-2000; 2000US-0220964.							
PR	PR	14-AUG-2000; 2000US-0224518.							
PR	PR	14-AUG-2000; 2000US-0224519.							
PR	PR	14-AUG-2000; 2000US-0225213.							
PR	PR	14-AUG-2000; 2000US-0225214.							
PR	PR	14-AUG-2000; 2000US-0225266.							
PR	PR	14-AUG-2000; 2000US-0225267.							
PR	PR	14-AUG-2000; 2000US-0225268.							
PR	PR	14-AUG-2000; 2000US-0225270.							
PR	PR	14-AUG-2000; 2000US-0225447.							
PR	PR	14-AUG-2000; 2000US-0225757.							
PR	PR	14-AUG-2000; 2000US-0225758.							
PR	PR	14-AUG-2000; 2000US-0225759.							
PR	PR	18-AUG-2000; 2000US-0226279.							
PR	PR	22-AUG-2000; 2000US-0226681.							
PR	PR	22-AUG-2000; 2000US-0226868.							
PR	PR	22-AUG-2000; 2000US-0227182.							
PR	PR	23-AUG-2000; 2000US-0227009.							
PR	PR	30-AUG-2000; 2000US-0228924.							
PR	PR	01-SEP-2000; 2000US-0229287.							
PR	PR	01-SEP-2000; 2000US-0229343.							
PR	PR	01-SEP-2000; 2000US-0229344.							
PR	PR	01-SEP-2000; 2000US-0229345.							

05-SEP-2000; 2000US-0229509.  
05-SEP-2000; 2000US-0229513.  
06-SEP-2000; 2000US-0230437.  
06-SEP-2000; 2000US-0230438.  
08-SEP-2000; 2000US-0231242.  
08-SEP-2000; 2000US-0231243.  
08-SEP-2000; 2000US-0231244.  
08-SEP-2000; 2000US-0231413.  
08-SEP-2000; 2000US-0231414.  
08-SEP-2000; 2000US-0232080.  
08-SEP-2000; 2000US-0232081.  
12-SEP-2000; 2000US-0231968.  
14-SEP-2000; 2000US-0232397.  
14-SEP-2000; 2000US-0232398.  
14-SEP-2000; 2000US-0232399.  
14-SEP-2000; 2000US-0232400.  
14-SEP-2000; 2000US-0232401.  
14-SEP-2000; 2000US-0233063.  
14-SEP-2000; 2000US-0233064.  
14-SEP-2000; 2000US-0233065.  
21-SEP-2000; 2000US-0234223.  
21-SEP-2000; 2000US-0234274.  
25-SEP-2000; 2000US-0234997.  
25-SEP-2000; 2000US-0234998.  
26-SEP-2000; 2000US-0235484.  
27-SEP-2000; 2000US-0235834.  
27-SEP-2000; 2000US-0235836.  
29-SEP-2000; 2000US-0236327.  
29-SEP-2000; 2000US-0236367.  
29-SEP-2000; 2000US-0236368.  
29-SEP-2000; 2000US-0236369.  
29-SEP-2000; 2000US-0236370.  
02-OCT-2000; 2000US-0236802.  
02-OCT-2000; 2000US-0237037.  
02-OCT-2000; 2000US-0237038.  
02-OCT-2000; 2000US-0237039.  
02-OCT-2000; 2000US-0237040.  
13-OCT-2000; 2000US-0239935.  
13-OCT-2000; 2000US-0239937.  
20-OCT-2000; 2000US-0240960.  
20-OCT-2000; 2000US-0241785.  
20-OCT-2000; 2000US-0241786.  
20-OCT-2000; 2000US-0241787.  
20-OCT-2000; 2000US-0241808.  
20-OCT-2000; 2000US-0241809.  
20-OCT-2000; 2000US-0241826.  
20-OCT-2000; 2000US-0242221.  
01-NOV-2000; 2000US-0244617.  
08-NOV-2000; 2000US-0246474.  
08-NOV-2000; 2000US-0246475.  
08-NOV-2000; 2000US-0246476.  
08-NOV-2000; 2000US-0246477.  
08-NOV-2000; 2000US-0246478.  
08-NOV-2000; 2000US-0246523.  
08-NOV-2000; 2000US-0246524.  
08-NOV-2000; 2000US-0246525.  
08-NOV-2000; 2000US-0246526.  
08-NOV-2000; 2000US-0246527.  
08-NOV-2000; 2000US-0246528.  
08-NOV-2000; 2000US-0246532.  
08-NOV-2000; 2000US-0246609.  
08-NOV-2000; 2000US-0246610.  
08-NOV-2000; 2000US-0246611.  
08-NOV-2000; 2000US-0246613.  
17-NOV-2000; 2000US-0249207.  
17-NOV-2000; 2000US-0249208.  
17-NOV-2000; 2000US-0249209.  
17-NOV-2000; 2000US-0249210.  
17-NOV-2000; 2000US-0249211.  
17-NOV-2000; 2000US-0249212.  
17-NOV-2000; 2000US-0249213.  
17-NOV-2000; 2000US-0249214.  
17-NOV-2000; 2000US-0249215.

PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250391.  
PR 01-DEC-2000; 2000US-0251160.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
PI WPI; 2001-541565/60.  
XX P-PSDB; ABB17948.  
DR  
DR Nucleic acids encoding 3224 human nervous system antigen polypeptides,  
XX useful for preventing, diagnosing and/or treating nervous system  
XX cancers and metastases -  
XX  
XX Claim 1; SEQ ID NO 3281; 1701pp + Sequence Listing; English.  
XX  
XX The invention relates to novel genes (ABA11004-ABA21534) and proteins  
XX (ABA14578-ABA18001) useful for preventing, treating or ameliorating  
XX medical conditions e.g. by protein or gene therapy. The genes are  
XX isolated from a range of human tissues disclosed in the specification.  
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful  
XX in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
XX and ovarian cancer and other cancers of the adrenal gland, bone, bone  
XX marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
XX (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
XX haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
XX disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
XX colitis; (c) cardiovascular disorders such as myocardial ischaemias;  
XX (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
XX epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
XX and parasitic infections.  
XX Note: The sequence data for this patent did not form part of the  
XX printed specification, but was obtained in electronic format directly  
XX from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 349 BP; 98 A; 96 C; 76 G; 76 T; 3 other;  
Query Match 58.5%; Score 34; DB 22; Length 349;  
Best Local Similarity 75.0%; Pred. No. 3.6;  
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
QY 2 CAGTGTGCTGAGAGCTGACATACCAGGACTGCTGAGAC 41  
Db 164 CAGTACTGGAGAGCTGACATACCAGGACTGCTGAGAC 125  
RESULT 13  
AAI91318/c  
ID AAI91318 standard; cDNA; 397 BP.  
XX  
XX AAI91318;  
XX

Human; ss; serine protease 8; cytostatic; virucidal; immunomodulatory; antiinflammatory; haemostatic; malignant tumour; haemopathy; HIV; human immunodeficiency virus infection; immunological disease; inflammation.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524
--	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

XX	OS	XX	FH	FT	FT	FT	PN	XX	PD	XX	PF	XX	PR	XX	PA	XX	PI	XX	DR	DR	XX	PT	PT	PT	PT	XX	PS	XX	CC	CC	CC	CC	CC	CC	CC	CC	CC	XX	SQ		Oy	Dd	RE	AA	ID	XX	AC	XX	DT	XX	DE	XX	KW	KW	XX	OS	XX	PN	XX	PD	XX	PF	XX	PR
----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	--	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----



23-AUG-2000; 2000US-0649167.

(HYSE-) HYSEQ INC.

Dmanac RT, Liu C, Tang YT;

WPI; 2001-639362/73.

P-PSDB; ABG10898.

New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity -

Claim 1; SEQ ID No 10889; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AA864197-AA894564 represent novel human diagnostic coding sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 992 BP; 226 A; 295 C; 276 G; 195 T; 0 other;

Query Match 58.5%; Score 24; DB 23; Length 992;

Best Local Similarity 75.0%; Pred. No. 4.5;

Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

2 CAGTTGCTGAGAGCTGTGACATACCGAGGACTGCTGAGAC 41

|||||

315 CAGTACTGGGAGGCTGACATGGTAGGATTGCTTGAGCC 276

arch completed: July 22, 2003, 12:24:07

b time : 174.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

nucleic - nucleic search, using sw model

on: July 22, 2003, 11:16:51 ; Search time 36 Seconds  
(without alignments)  
349.271 Million cell updates/sec

le: U95626-C-AT-42723

fect score: 41

quence: 1 tcagtgtctgagactga.....ataccagactgcctgagac 41

ring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

rchd: 441362 seqs, 1533381 residues

al number of hits satisfying chosen parameters: 882724

imum DB seq length: 0

ximum DB seq length: 2000000000

st-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

abase : Issued Patents NA:\*

1: /cgn2\_6/prodata/1/ina/5A COMB.seq:\*

2: /cgn2\_6/prodata/1/ina/5B COMB.seq:\*

3: /cgn2\_6/prodata/1/ina/6A COMB.seq:\*

4: /cgn2\_6/prodata/1/ina/6B COMB.seq:\*

5: /cgn2\_6/prodata/1/ina/PTUS COMB.seq:\*

6: /cgn2\_6/prodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

sult No.	Score	Query Match %	Length	DB ID	Description
1	24	58.5	3565	1	US-08-578-649-3
2	22.4	54.6	5892	3	US-08-578-649-3
3	22.4	54.6	15328	2	US-08-888-497-33
4	22.4	54.6	15328	4	US-09-362-230-33
5	22.4	54.6	15328	5	PCT-US94-07926-33
6	22.4	54.6	87350	3	US-08-781-891-79
7	22.4	54.6	87543	4	US-09-791-211-3
8	22	53.7	5162	4	US-08-916-917-13
9	22	53.7	5162	3	US-09-225-170-13
10	21.8	53.2	393	4	US-09-574-141A-97
11	21.8	53.2	399	4	US-09-574-141A-94
12	21.8	53.2	907	3	US-09-081-320-39
13	21.8	53.2	907	4	US-09-574-141A-39
14	21.8	53.2	907	4	US-09-707-780-39
15	21.8	53.2	3224	4	US-08-965-729A-2
16	21.8	53.2	6485	3	US-09-320-2
17	21.8	53.2	6485	4	US-09-574-141A-2
18	21.8	53.2	6485	4	US-09-707-780-2
19	21.8	53.2	8743	3	US-09-081-320-1
20	21.8	53.2	8743	3	US-09-574-141A-1
21	21.8	53.2	8743	3	US-09-707-780-1
22	21.4	52.2	1866	4	US-08-909-742-1
23	21.4	52.2	1866	4	US-09-412-289-1
24	21.4	52.2	43950	4	US-09-735-934A-3
25	21.2	51.7	12141	4	US-09-488-671-10
26	21	51.2	1845	4	US-08-887-534A-22
27	20.8	50.7	75	4	US-09-461-697-151

Sequence 91, Appl  
Sequence 126, Appl  
Sequence 33, Appl  
Sequence 35, Appl  
Sequence 84, Appl  
Sequence 57, Appl  
Sequence 84, Appl  
Sequence 84, Appl  
Sequence 40, Appl  
Sequence 10, Appl  
Sequence 1, Appl  
Sequence 44, Appl  
Sequence 1, Appl  
Sequence 17, Appl  
Sequence 90, Appl

ALIGNMENTS

RESULT 1  
US-08-578-649-3  
; Sequence 3, Application US/08578649  
; Patent No. 5770366  
; GENERAL INFORMATION:  
; APPLICANT: Ulrich Bogdan  
; APPLICANT: Reinhard Buttner  
; APPLICANT: Brigitte Kaluza  
; TITLE OF INVENTION: MELANOMA-INHIBITING PROTEIN  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Felfe & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/578,649  
; FILING DATE: 29-July-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE P 43 24 247.2  
; FILING DATE: 20-July-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Andrew L. Tiajolloff  
; REGISTRATION NUMBER: 31,575  
; REFERENCE/DOCKET NUMBER: BOER 1035-PFF/ALT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 688-9200  
; TELEFAX: (212) 838-3884  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3565 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: sig peptide  
; LOCATION: 1378..1449  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 1378..1504  
; FEATURE:

NAME/KEY: exon  
LOCATION: 1586..1719  
FEATURE:  
NAME/KEY: exon  
LOCATION: 2804..2914  
FEATURE:  
NAME/KEY: exon  
LOCATION: 3232..3252  
FEATURE:  
NAME/KEY: -  
LOCATION: one-of(2216)  
OTHER INFORMATION: /note= "N in position 2216  
OTHER INFORMATION: denotes an indefinite number and sequence of  
OTHER INFORMATION: nucleotides"  
-08-578-649-3

Query Match 58.5%; Score 24; DB 1; Length 3565;  
Best Local Similarity 75.0%; Pred. No. 0.49;  
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

2 CAGTTGCTGAGAGCTGCATACACGAGCTGCTGAGAC 41  
|||||  
519 CAGTACTCAGGAGCTGAGGTGGAGGATGCTGAGTC 558

SULT 2  
-08-755-587-27/c  
Sequence 27, Application US/08755587  
Patent No. 6045997  
GENERAL INFORMATION:  
APPLICANT: Futreal, Phillip A  
APPLICANT: Wooster, Richard F  
APPLICANT: Ashworth, Alan  
APPLICANT: Stratton, Michael R  
TITLE OF INVENTION: Materials and methods relating to the  
TITLE OF INVENTION: identification and sequencing of the BRCA2 cancer  
NUMBER OF SEQUENCES: 222  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Bell Seltzer Park & Gibson  
STREET: 310 UCB Plaza, 3605 Glenwood Avenue, PO Drawer 31107  
CITY: Raleigh  
STATE: NC  
COUNTRY: USA

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/755,587  
FILING DATE: 25-NOV-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9523959.6  
FILING DATE: 23-NOV-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9525555.0  
FILING DATE: 14-DEC-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9617961.9  
FILING DATE: 28-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Kenneth D Sibley  
REGISTRATION NUMBER: 31,665  
REFERENCE/DOCKET NUMBER: 5405-135  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5892 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:

NAME/KEY: CDS  
LOCATION: 483..5412  
FEATURE:  
NAME/KEY: exon  
LOCATION: 481..5412  
US-08-755-587-27

Query Match 54.6%; Score 22.4; DB 3; Length 5892;  
Best Local Similarity 72.5%; Pred. No. 2.7;  
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 CAGTTGCTGAGAGCTGCATACACGAGCTGCTGAGAC 41  
|||||  
Db 231 CAGTACTCAGGAGCTGAGGTGCAAGACTGCTGAGCC 192

RESULT 3  
US-08-888-497-33  
Sequence 33, Application US/08888497  
Patent No. 5972677  
GENERAL INFORMATION:  
APPLICANT: Seilhamer, Jeffrey J.  
TITLE OF INVENTION: Mammalian Phospholipase A2 Nucleotide  
TITLE OF INVENTION: Sequences and Low Molecular Weight Amino Acid Sequences  
TITLE OF INVENTION: Encoded Thereby, Antisense Sequences and Nucleotide  
TITLE OF INVENTION: Sequences Having Internal Ribosome Binding Sites  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ruden, Barnett, McClosky, Smith, Schuster &  
ADDRESSEE: Russell PA  
STREET: 200 East Broward Boulevard  
CITY: Fort Lauderdale  
STATE: FL  
COUNTRY: USA  
ZIP: 33301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/888,497  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/651,405  
FILING DATE:  
APPLICATION NUMBER: US 08/097,354  
FILING DATE: 26-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Manso, Peter J.  
REGISTRATION NUMBER: 32,264  
REFERENCE/DOCKET NUMBER: IN21044-5  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 305-527-2498  
TELEFAX: 305-764-4996  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15328 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-888-497-33

Query Match 54.6%; Score 22.4; DB 2; Length 15328;  
Best Local Similarity 72.5%; Pred. No. 3.4;  
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 CAGTTGCTGAGAGCTGCATACACGAGCTGCTGAGAC 41  
|||||  
Db 8735 CAGTACTCAGAGGCTGAGATAGGAGATCACTGAGCC 8774

SULT 4  
-09-362-230-33  
Sequence 33, Application US/09362230  
Patent No. 6352849  
GENERAL INFORMATION:  
APPLICANT: Tischfield, Jay A.  
APPLICANT: Seilhamer, Jeffrey J.  
TITLE OF INVENTION: Mammalian Phospholipase A2 Nucleotide  
TITLE OF INVENTION: Sequences and Low Molecular Weight Amino Acid Sequences  
TITLE OF INVENTION: Encoded Thereby, Antisense Sequences and Nucleotide  
TITLE OF INVENTION: Sequences Having Internal Ribosome Binding Sites  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ruden, Barnett, McClosky, Smith, Schuster &  
ADDRESSEE: Russell PA  
STREET: 200 East Broward Boulevard  
CITY: Fort Lauderdale  
STATE: FL  
COUNTRY: USA  
ZIP: 33301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/362,230  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/888,497  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Manso, Peter J.  
REGISTRATION NUMBER: 32,264  
REFERENCE/DOCKET NUMBER: IN21044-5  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 305-527-2498  
TELEFAX: 305-764-4996  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15328 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
-09-362-230-33  
Query Match 54.6%; Score 22.4; DB 4; Length 15328;  
Best Local Similarity 72.5%; Pred. No. 3.4;  
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
2 CAGTTGCTGAGAGGCTGACATACAGGACTGCCTGAGAC 41  
|||||  
8735 CAGCTACTGAGAGGCTGAGATAGGAGGATCAGCTGAGCC 8774  
SULT 5  
1-US94-07926-33  
Sequence 33, Application PC/TUS9407926  
GENERAL INFORMATION:  
APPLICANT: Tischfield, Jay A.  
APPLICANT: Seilhamer, Jeffrey J.  
TITLE OF INVENTION: Mammalian Phospholipase A2 Nucleotide  
TITLE OF INVENTION: Sequences and Low Molecular Weight Amino Acid Sequences  
TITLE OF INVENTION: Encoded Thereby, Antisense Sequences and Nucleotide  
TITLE OF INVENTION: Sequences Having Internal Ribosome Binding Sites  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ruden, Barnett, McClosky, Smith, Schuster &  
ADDRESSEE: Russell PA  
STREET: 200 East Broward Boulevard  
CITY: Fort Lauderdale  
STATE: FL  
COUNTRY: USA  
ZIP: 33301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/07926  
FILING DATE: 15-JUL-1994  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/097,354  
FILING DATE: 26-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Manso, Peter J.  
REGISTRATION NUMBER: 32,264  
REFERENCE/DOCKET NUMBER: IN21044-5  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 305-527-2498  
TELEFAX: 305-764-4996  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15328 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
-09-362-230-33  
Query Match 54.6%; Score 22.4; DB 4; Length 15328;  
Best Local Similarity 72.5%; Pred. No. 3.4;  
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
2 CAGTTGCTGAGAGGCTGACATACAGGACTGCCTGAGAC 41  
|||||  
8735 CAGCTACTGAGAGGCTGAGATAGGAGGATCAGCTGAGCC 8774

ADDRESSEE: Ruden, Barnett, McClosky, Smith, Schuster &  
ADDRESSEE: Russell PA  
STREET: 200 East Broward Boulevard  
CITY: Fort Lauderdale  
STATE: FL  
COUNTRY: USA  
ZIP: 33301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/07926  
FILING DATE: 15-JUL-1994  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/097,354  
FILING DATE: 26-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Manso, Peter J.  
REGISTRATION NUMBER: 32,264  
REFERENCE/DOCKET NUMBER: IN21044-5  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 305-527-2498  
TELEFAX: 305-764-4996  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15328 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
PCT-US94-07926-33  
Query Match 54.6%; Score 22.4; DB 5; Length 15328;  
Best Local Similarity 72.5%; Pred. No. 3.4;  
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
2 CAGTTGCTGAGAGGCTGACATACAGGACTGCCTGAGAC 41  
|||||  
8735 CAGCTACTGAGAGGCTGAGATAGGAGGATCAGCTGAGCC 8774  
RESULT 6  
US-08-781-891-79/c  
; Sequence 79, Application US/08781891  
; Patent No. 6090620  
; GENERAL INFORMATION:  
; APPLICANT: Fu, Ying-Hui  
; APPLICANT: Yu, Chang-En  
; APPLICANT: Oshima, Junko  
; APPLICANT: Mulligan, John T.  
; APPLICANT: Schellenberg, Gerald D.  
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO  
; TITLE OF INVENTION: WERNER'S SYNDROME  
; NUMBER OF SEQUENCES: 209  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED AND BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/781,891  
; FILING DATE: 27-DEC-1996  
; CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:  
NAME: No. 6090620tenburg Ph.D., Carol  
REGISTRATION NUMBER: 39,317  
REFERENCE/DOCKET NUMBER: 240052.419  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 79:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 87350 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
-08-781-891-79

Query Match 54.6%; Score 22.4; DB 3; Length 87350;  
Best Local Similarity 72.5%; Pred.No.5.5;  
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

2 CAGTTGCTGAGAGCCTGACATACCAGGACTGCTGAGAC 41  
|||||  
63492 CAGTACTCAGGATGCTGAGATPAGAGGACCTGTTGAGCC 63453

SULT 7  
-09-791-211-3/c  
Sequence 3, Application US/09791211  
Patent No. 6448080  
GENERAL INFORMATION:  
APPLICANT: Donna T. Ward  
APPLICANT: Andrew T. Watt  
TITLE OF INVENTION: ANTISENSE MODULATION OF WRN EXPRESSION  
FILE REFERENCE: RTS-0205  
CURRENT APPLICATION NUMBER: US/09/791,211  
CURRENT FILING DATE: 2001-02-23  
NUMBER OF SEQ ID NOS: 90  
SEQ ID NO 3  
LENGTH: 87543  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 7421  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 7427  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 11609  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 12605  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 12742  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 29370  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 29422  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 29979  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 29980  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 29981  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 30136  
OTHER INFORMATION: unknown

OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 30140  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 31205  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 31206  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 31592  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 33095  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 33160  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 34066  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 34072  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 36816  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 39020  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 42164  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 42459  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 46808  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 46823  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 46826  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 47291  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 52786  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 52787  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 53384  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 54684  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 59215  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 59235  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 59242  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 63290  
OTHER INFORMATION: unknown

NAME/KEY: unsure  
LOCATION: 66614  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 68660  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 68697  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 68718  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 68733  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 68739  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 69785  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 79134  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 79198  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 86336  
OTHER INFORMATION: unknown  
OTHER INFORMATION:  
-09-791-211-3

Query Match 54.6%; Score 22.4; DB 4; Length 87543;  
Best Local Similarity 72.5%; Pred. No. 5.5;  
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

2 CAGTTGCTGAGAGCTGACATACAGGACTGCTGAGAC 41  
63695 CAGTACTAGGATGCTGATAGAGAGCTGCTGAGCC 63646

SULT 8  
-08-916-917-13  
Sequence 13, Application US/08916917  
Patent No. 5856132  
GENERAL INFORMATION:  
APPLICANT: Stephens, Len  
APPLICANT: Hawkins, Phillip Thomas  
APPLICANT: Braselmann, Sylvia  
TITLE OF INVENTION: G-BETA-GAMMA REGULATED  
TITLE OF INVENTION: PHOSPHATIDYLINOSITOL-3', KINASE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds, LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036-2811  
COMPUTER READABLE FORM:  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: FASTSEQ for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/916,917  
FILING DATE: 15-AUG-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/672,211  
FILING DATE: 27-JUN-1996  
ATTORNEY/AGENT INFORMATION:

NAME: Abrams, Samuel B  
REGISTRATION NUMBER: 30,605  
REFERENCE/DOCKET NUMBER: 8549-0006-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-493-4935  
TELEFAX: 650-493-5556  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5162 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-916-917-13

Query Match 53.7%; Score 22; DB 2; Length 5162;  
Best Local Similarity 70.0%; Pred. No. 3.8;  
Matches 28; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

Qy 2 CAGTTGCTGAGAGCTGACATACAGGACTGCTGAGAC 41  
Db 5012 CAGTACTAGGAGCTGAGCAGGAGGATTGCTGAGCC 5051

RESULT 9  
US-09-225-170-13  
Sequence 13, Application US/09225170  
Patent No. 6017763  
GENERAL INFORMATION:  
APPLICANT: Stephens, Len  
APPLICANT: Hawkins, Phillip Thomas  
APPLICANT: Braselmann, Sylvia  
TITLE OF INVENTION: G-BETA-GAMMA REGULATED  
TITLE OF INVENTION: PHOSPHATIDYLINOSITOL-3', KINASE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds, LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036-2811  
COMPUTER READABLE FORM:  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: FASTSEQ for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/225,170  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/916,917  
FILING DATE: 15-AUG-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Abrams, Samuel B  
REGISTRATION NUMBER: 30,605  
REFERENCE/DOCKET NUMBER: 8549-0006-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-493-4935  
TELEFAX: 650-493-5556  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5162 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-225-170-13

Query Match 53.7%; Score 22; DB 3; Length 5162;  
Best Local Similarity 70.0%; Pred. No. 3.8;  
Matches 28; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

2 CAGTTGCTGAGAGCTGCATACACAGAGCTGCCTGAGAC 41  
|||||  
5012 CAGGTACTCAGAGGCTGAGCGAGGATTCCTGAGCC 5051

SULT 10  
-09-574-141A-97  
Sequence 97, Application US/09574141A  
Patent No. 6395490

GENERAL INFORMATION:  
APPLICANT: Gonsalves, Dennis  
APPLICANT: Meng, Baozhong  
TITLE OF INVENTION: RUPESTRIS STEM PITCHING ASSOCIATED VIRUS  
TITLE OF INVENTION: NUCLEIC ACIDS, PROTEINS, AND THEIR USES  
FILE REFERENCE: 07678/035005  
CURRENT APPLICATION NUMBER: US/09/574,141A  
CURRENT FILING DATE: 2000-05-18  
PRIOR FILING DATE: 1997-05-20  
PRIOR FILING DATE: 1997-05-20  
PRIOR FILING DATE: 1997-12-17  
PRIOR APPLICATION NUMBER: 60/069,902  
PRIOR APPLICATION NUMBER: 09/081,320  
NUMBER OF SEQ ID NOS: 97  
SOFTWARE: FastSEQ for Windows Version 4.0  
SEQ ID NO 97  
LENGTH: 393  
TYPE: DNA  
ORGANISM: Rupestris stem pitting associated virus  
-09-574-141A-97

Query Match 53.2%; Score 21.8; DB 4; Length 393;  
Best Local Similarity 78.8%; Pred. No. 2.3;  
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

7 GCTGAGAGCTGCATACACAGAGCTGCCTGAG 39  
|||||  
166 GCTGAGCAGAGTGCATACCTTAGACTGCTTGAG 198

SULT 11  
-09-574-141A-94  
Sequence 94, Application US/09574141A  
Patent No. 6395490

GENERAL INFORMATION:  
APPLICANT: Gonsalves, Dennis  
APPLICANT: Meng, Baozhong  
TITLE OF INVENTION: RUPESTRIS STEM PITCHING ASSOCIATED VIRUS  
TITLE OF INVENTION: NUCLEIC ACIDS, PROTEINS, AND THEIR USES  
FILE REFERENCE: 07678/035005  
CURRENT APPLICATION NUMBER: US/09/574,141A  
CURRENT FILING DATE: 2000-05-18  
PRIOR FILING DATE: 1997-05-20  
PRIOR FILING DATE: 1997-05-20  
PRIOR FILING DATE: 1997-12-17  
PRIOR FILING DATE: 1997-12-17  
PRIOR APPLICATION NUMBER: 09/081,320  
PRIOR FILING DATE: 1998-05-19  
NUMBER OF SEQ ID NOS: 97  
SOFTWARE: FastSEQ for Windows Version 4.0  
SEQ ID NO 94  
LENGTH: 399  
TYPE: DNA  
ORGANISM: Rupestris stem pitting associated virus  
-09-574-141A-94

Query Match 53.2%; Score 21.8; DB 4; Length 399;  
Best Local Similarity 78.8%; Pred. No. 2.3;  
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

7 GCTGAGAGCTGCATACACAGAGCTGCCTGAG 39  
|||||

Db 169 GCTGAGCAGAGTGACATACCTTAGACTGCTTGAG 201

## RESULT 12

US-09-081-320-39/c  
Sequence 39, Application US/09081320  
Patent No. 6093544  
GENERAL INFORMATION:  
APPLICANT: Gonsalves, Dennis  
APPLICANT: Meng, Baozhong  
TITLE OF INVENTION: RUPESTRIS STEM PITCHING ASSOCIATED VIRUS  
TITLE OF INVENTION: NUCLEIC ACIDS, PROTEINS, AND THEIR USES  
NUMBER OF SEQUENCES: 54  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP  
STREET: Clinton Square, P.O. Box 1051  
CITY: Rochester  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 14603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/081,320  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/047,147  
FILING DATE: 20-MAY-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/069,902  
FILING DATE: 17-DEC-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Goldman, Michael L.  
REGISTRATION NUMBER: 30,727  
REFERENCE/DOCKET NUMBER: 19603/1722  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (716) 263-1304  
TELEFAX: (716) 263-1600  
INFORMATION FOR SEQ ID NO: 39:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 907 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cdna  
US-09-081-320-39

Query Match 53.2%; Score 21.8; DB 3; Length 907;  
Best Local Similarity 78.8%; Pred. No. 2.9;  
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 7 GCTGAGAGCTGCATACACAGAGCTGCCTGAG 39  
|||||

Db 410 GCTGAGCAGAGTGACATACCTTAGACTGCTTGAG 378

## RESULT 13

US-09-574-141A-39/c  
Sequence 39, Application US/09574141A  
Patent No. 6395490  
GENERAL INFORMATION:  
APPLICANT: Gonsalves, Dennis  
APPLICANT: Meng, Baozhong  
TITLE OF INVENTION: RUPESTRIS STEM PITCHING ASSOCIATED VIRUS  
TITLE OF INVENTION: NUCLEIC ACIDS, PROTEINS, AND THEIR USES  
FILE REFERENCE: 07678/035005  
CURRENT APPLICATION NUMBER: US/09/574,141A  
CURRENT FILING DATE: 2000-05-18  
PRIOR APPLICATION NUMBER: 60/047,147

```
PRIOR FILING DATE: 1997-05-20
PRIOR APPLICATION NUMBER: 60/069,902
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 09/081,320
PRIOR FILING DATE: 1998-05-19
NUMBER OF SEQ ID NOS: 97
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 39
LENGTH: 907
TYPE: DNA
ORGANISM: Rupestris stem pitting associated virus
09-574-141A-39

Query Match 53.2%; Score 21.8; DB 4; Length 907;
Best Local Similarity 78.8%; Pred. No. 2.9;
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

7 GCTGAGAGCCTGACATACACGAGCTCCTTGAG 39
||||| ||| ||||| ||||| ||||| |||||
410 GCTGACAGAGTGACATCTTAGACTCTTGAG 378

ULT 14
09-707-780-39/c
Sequence 39, Application US/09707780
Patent No. 639308
GENERAL INFORMATION:
APPLICANT: Gonsalves, Dennis
APPLICANT: Meng, Baozhong
TITLE OF INVENTION: RUPESTRIS STEM PITTING ASSOCIATED VIRUS
TITLE OF INVENTION: NUCLEIC ACIDS, PROTEINS, AND THEIR USES
FILE REFERENCE: 07678/035006
CURRENT APPLICATION NUMBER: US/09/707,780
CURRENT FILING DATE: 2000-11-07
PRIOR APPLICATION NUMBER: 09/081,320
PRIOR FILING DATE: 1998-05-19
PRIOR APPLICATION NUMBER: 60/047,147
PRIOR FILING DATE: 1997-05-20
PRIOR APPLICATION NUMBER: 60/069,902
PRIOR FILING DATE: 1997-12-17
NUMBER OF SEQ ID NOS: 54
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 39
LENGTH: 907
TYPE: DNA
ORGANISM: Rupestris stem pitting associated virus
09-707-780-39

Query Match 53.2%; Score 21.8; DB 4; Length 907;
Best Local Similarity 78.8%; Pred. No. 2.9;
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

7 GCTGAGAGCCTGACATACACGAGCTCCTTGAG 39
||||| ||| ||||| ||||| ||||| |||||
410 GCTGACAGAGTGACATCTTAGACTCTTGAG 378

ULT 15
08-965-729A-2/c
Sequence 2, Application US/08965729A
Patent No. 6200751
GENERAL INFORMATION:
APPLICANT: Jian-Ming Gu and Charles T. Eamon
TITLE OF INVENTION: ENOTHELUM SPECIFIC EXPRESSION
TITLE OF INVENTION: REGULATED BY EPCR CONTROL ELEMENTS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center, 1201 West Peachtree Street
CITY: Atlanta
STATE: GA
COUNTRY: USA
ZIP: 30309-4530
```

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/965,729A
FILING DATE: 07-NOV-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/030,718
FILING DATE: 08-NOV-1997
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMRF 164 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-873-8794
TELEFAX: 404-873-8795
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3224 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: /note= "Nucleotides 2272 through 2729 are a
OTHER INFORMATION: thrombin responsive element"; Human
US-08-965-729A-2

Query Match 53.2%; Score 21.8; DB 4; Length 3224;
Best Local Similarity 70.7%; Pred. No. 4.1;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 TCAGTTGCTGAGAGCTTGACATACACGAGCTCCTTGAGAC 41
||||| ||| ||||| ||||| ||||| |||||
DB 198 TCAGTACTCAGGAGCTGAGTAGGAGGATTGCTTGAGCC 158

Search completed: July 22, 2003, 11:18:22
Job time : 38 secs
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

nucleic - nucleic search, using sw model

on: July 22, 2003, 11:16:53 ; Search time 130 Seconds  
(without alignments)  
650.639 Million cell updates/sec

le: U95626-C-AT-42723

ence: 1 tcagtgtcgaagactga.....ataccagactgctgagac 41

ring table: IDENTIFY\_NUC  
Gapop 10.0 , Gapext 1.0

rchd: 1439767 seqs, 1031500376 residues

al number of hits satisfying chosen parameters: 2879534

imum DB seq length: 0

imum DB seq length: 2000000000

t-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

abase : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq1.\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq2.\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq3.\*
- 13: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

ult	No.	Score	Query Match	Length	ID	Description
1	41	100.0	143068	11	US-09-967-768A-316	Sequence 316, App
2	25.2	61.5	591	15	US-10-198-846-11772	Sequence 11772, A
3	24	58.5	544	15	US-10-027-632-216678	Sequence 216678, A
4	24	58.5	544	15	US-10-027-632-152011	Sequence 152011, A
5	24	58.5	544	15	US-10-027-632-103171	Sequence 103171, A
6	24	58.5	544	15	US-10-027-632-103172	Sequence 103172, A
7	24	58.5	544	15	US-10-027-632-224761	Sequence 224761, A
8	24	58.5	544	15	US-10-027-632-224761	Sequence 224761, A
9	24	58.5	552	15	US-10-027-632-224761	Sequence 224761, A
10	24	58.5	573	15	US-10-027-632-224761	Sequence 224761, A
11	24	58.5	573	15	US-10-027-632-224761	Sequence 224761, A
12	24	58.5	17335	11	US-09-764-847-1280	Sequence 1280, Ap
13	24	58.5	17335	11	US-10-092-154-1280	Sequence 1280, Ap
14	24	58.5	19882	11	US-09-764-847-1281	Sequence 1281, Ap
15	24	58.5	19882	15	US-10-092-154-1281	Sequence 1281, Ap
16	24	58.5	183337	15	US-10-020-141-5	Sequence 5, Appli

Sequence 3, Appli  
Sequence 3, Appli  
Sequence 33, Appli  
Sequence 216678,  
Sequence 152011,  
Sequence 103171,  
Sequence 103172,  
Sequence 3463, Ap  
Sequence 224761,  
Sequence 14311,  
Sequence 5856, A  
Sequence 301953,  
Sequence 222706,  
Sequence 222707,  
Sequence 15461, A  
Sequence 145792,  
Sequence 2178, Ap  
Sequence 214260,  
Sequence 273704,  
Sequence 273705,  
Sequence 319580,  
Sequence 319581,  
Sequence 319582,  
Sequence 86169, A  
Sequence 206768,  
Sequence 239374,  
Sequence 256903,  
Sequence 244369,

ALIGNMENTS

RESULT 1

US-09-967-768A-316  
; Sequence 316, Application US/09967768A

; Patent No. US20020150877A1

; GENERAL INFORMATION:

; APPLICANT: Augustus, Meena

; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu

; FILE REFERENCE: 689290-72

; CURRENT APPLICATION NUMBER: US/09/967,768A

; CURRENT FILING DATE: 2001-09-28

; PRIOR APPLICATION NUMBER: US/60/236,109

; PRIOR FILING DATE: 2000-09-28

; PRIOR APPLICATION NUMBER: US/60/236,034

; PRIOR FILING DATE: 2000-09-28

; PRIOR APPLICATION NUMBER: US/60/236,111

; PRIOR FILING DATE: 2000-09-28

; NUMBER OF SEQ ID NOS: 325

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 316

; LENGTH: 143068

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-967-768A-316

Query Match 100.0%; Score 41; DB 11; Length 143068;  
Best Local Similarity 100.0%; Pred. No. 5.2e-07;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCAGTTCCTGAGAGCTGACATACAGGACTGCTGAGAC 41

Db 42703 TCAGTTCCTGAGAGCTGACATACAGGACTGCTGAGAC 42743

RESULT 2

US-10-198-846-11772  
; Sequence 11772, Application US/10198846

; Publication No. US2003009974A1

; GENERAL INFORMATION:

APPLICANT: Lillie, James  
APPLICANT: Xu, Yongyao  
APPLICANT: Wang, Youzhen  
APPLICANT: Steinmann, Kathleen  
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS  
TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
TITLE OF INVENTION: THERAPY OF BREAST CANCER  
FILE REFERENCE: MEI-049  
CURRENT APPLICATION NUMBER: US/10/198,846  
CURRENT FILING DATE: 2002-07-18  
PRIOR APPLICATION NUMBER: 60/306,220  
PRIOR FILING DATE: 2001-07-18  
NUMBER OF SEQ ID NOS: 14084  
SOFTWARE: FastSEQ for Windows Version 4.0  
SEQ ID NO 11772  
LENGTH: 591  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 576,577,578,579,580,581,582,583,584,585,586,587,  
LOCATION: 588,589,590,591  
OTHER INFORMATION: n = A,T,C or G  
-10-198-846-11772

Query Match 61.5%; Score 25.2; DB 15; Length 591;  
Best Local Similarity 78.9%; Pred. No. 0.86;  
Matches 30; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

1 TCAGTCTGAGAGCCTGACATACCGAGCTGCTGA 38  
491 TCAGTCTGAGAGCCTGACCTGAGTGGAGTGGTGA 528

SULT 3  
-10-027-632-66000  
Sequence 66000, Application US/10027632  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
TITLE OF INVENTION: Polymorphisms in the Human Genome  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSEQ for Windows Version 4.0  
SEQ ID NO 66000  
LENGTH: 544  
TYPE: DNA  
ORGANISM: Human  
-10-027-632-66000

Query Match 58.5%; Score 24; DB 15; Length 544;  
Best Local Similarity 75.0%; Pred. No. 2.6;  
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

2 CAGTTCGTGAGAGCCTGACATACCGAGCTGCTGAGAC 41  
127 CAGCTACTCAGGAGGCTGAGTGCAGGATTGCTTGAGAC 166

RESULT 4  
US-10-027-632-66001  
Sequence 66001, Application US/10027632  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
TITLE OF INVENTION: Polymorphisms in the Human Genome  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSEQ for Windows Version 4.0  
SEQ ID NO 66001  
LENGTH: 544  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-66001

Query Match 58.5%; Score 24; DB 15; Length 544;  
Best Local Similarity 75.0%; Pred. No. 2.6;  
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2 CAGTTCGTGAGAGCCTGACATACCGAGCTGCTGAGAC 41  
DB 127 CAGCTACTCAGGAGGCTGAGTGCAGGATTGCTTGAGAC 166

RESULT 5  
US-10-027-632-66002  
Sequence 66002, Application US/10027632  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
TITLE OF INVENTION: Polymorphisms in the Human Genome  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSEQ for Windows Version 4.0  
SEQ ID NO 66002  
LENGTH: 544  
TYPE: DNA  
ORGANISM: Human

10-027-632-66002

Query Match 58.5%; Score 24; DB 15; Length 544;

Best Local Similarity 75.0%; Pred. No. 2.6;

Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

2 CAGTTGCTGAGAAGCTGCATACACAGGACTGCTTGAGAC 41

127 CAGCTACTCAGGAGGCTGAGGTGCCAGGATTGCTTGAGAC 166

SULT 6

10-027-632-298997

Sequence 298997, Application US/10027632

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

FILE OF INVENTION: Polymorphisms in the Human Genome

FILE REFERENCE: 108827.129

CURRENT APPLICATION NUMBER: US/10/027,632

CURRENT FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US 60/218,006

PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/198,676

PRIOR FILING DATE: 2000-04-20

PRIOR APPLICATION NUMBER: US 60/193,483

PRIOR FILING DATE: 2000-03-29

PRIOR APPLICATION NUMBER: US 60/185,218

PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/167,363

PRIOR FILING DATE: 1999-11-23

PRIOR APPLICATION NUMBER: US 60/156,358

PRIOR FILING DATE: 1999-09-28

PRIOR APPLICATION NUMBER: US 60/146,002

PRIOR FILING DATE: 1999-08-09

NUMBER OF SEQ ID NOS: 325720

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 298997

LENGTH: 544

TYPE: DNA

ORGANISM: Human

10-027-632-298997

Query Match

Best Local Similarity 58.5%; Score 24; DB 15; Length 544;

Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

2 CAGTTGCTGAGAAGCTGCATACACAGGACTGCTTGAGAC 41

127 CAGCTACTCAGGAGGCTGAGGTGCCAGGATTGCTTGAGAC 166

SULT 7

10-027-632-298998

Sequence 298998, Application US/10027632

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

FILE OF INVENTION: Polymorphisms in the Human Genome

FILE REFERENCE: 108827.129

CURRENT APPLICATION NUMBER: US/10/027,632

CURRENT FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US 60/218,006

PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/198,676

PRIOR FILING DATE: 2000-04-20

PRIOR APPLICATION NUMBER: US 60/193,483

PRIOR FILING DATE: 2000-03-29

PRIOR APPLICATION NUMBER: US 60/185,218

PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/167,363

PRIOR FILING DATE: 1999-11-23

PRIOR APPLICATION NUMBER: US 60/156,358

PRIOR FILING DATE: 1999-09-28

PRIOR APPLICATION NUMBER: US 60/146,002

PRIOR FILING DATE: 1999-08-09

NUMBER OF SEQ ID NOS: 325720

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 298998

LENGTH: 544

TYPE: DNA

ORGANISM: Human

US-10-027-632-298998

Query Match

Best Local Similarity 58.5%; Score 24; DB 15; Length 544;

Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

2 CAGTTGCTGAGAAGCTGCATACACAGGACTGCTTGAGAC 41

127 CAGCTACTCAGGAGGCTGAGGTGCCAGGATTGCTTGAGAC 166

RESULT 8

US-10-027-632-298999

Sequence 298999, Application US/10027632

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

FILE OF INVENTION: Polymorphisms in the Human Genome

FILE REFERENCE: 108827.129

CURRENT APPLICATION NUMBER: US/10/027,632

CURRENT FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US 60/218,006

PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/198,676

PRIOR FILING DATE: 2000-04-20

PRIOR APPLICATION NUMBER: US 60/193,483

PRIOR FILING DATE: 2000-03-29

PRIOR APPLICATION NUMBER: US 60/185,218

PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/167,363

PRIOR FILING DATE: 1999-11-23

PRIOR APPLICATION NUMBER: US 60/156,358

PRIOR FILING DATE: 1999-09-28

PRIOR APPLICATION NUMBER: US 60/146,002

PRIOR FILING DATE: 1999-08-09

NUMBER OF SEQ ID NOS: 325720

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 298999

LENGTH: 544

TYPE: DNA

ORGANISM: Human

US-10-027-632-298999

Query Match

Best Local Similarity 58.5%; Score 24; DB 15; Length 544;

Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

2 CAGTTGCTGAGAAGCTGCATACACAGGACTGCTTGAGAC 41

127 CAGCTACTCAGGAGGCTGAGGTGCCAGGATTGCTTGAGAC 166

RESULT 9

US-10-027-632-286697/c

Sequence 286697, Application US/10027632

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

FILE OF INVENTION: Polymorphisms in the Human Genome

FILE REFERENCE: 108827.129

CURRENT APPLICATION NUMBER: US/10/027,632

CURRENT FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US 60/218,006

PRIOR FILING DATE: 2000-07-12

```

PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 286697
LENGTH: 552
TYPE: DNA
ORGANISM: Human
-10-027-632-286697

Query Match          58.5%; Score 24; DB 15; Length 552;
Best Local Similarity 75.0%; Pred. No. 2.6;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

      2 CAGTTCGTGAGAGCTGCATACACAGGACTGCTTGAGAC 41
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
300 CAGCTACTGGGAGGCTGCATGTTAGGATTGCTTGAGCC 261

SULT 10
-10-027-632-288310
Sequence 288310, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
TITLE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 288310
LENGTH: 573
TYPE: DNA
ORGANISM: Human
-10-027-632-288310

Query Match          58.5%; Score 24; DB 15; Length 573;
Best Local Similarity 75.0%; Pred. No. 2.6;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

      2 CAGTTCGTGAGAGCTGCATACACAGGACTGCTTGAGAC 41
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
270 CAGCTACTGGGAGGCTGCATGTTAGGATTGCTTGAGCC 309

SULT 11
-10-027-632-288311
Sequence 288311, Application US/10027632
```

```

GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
TITLE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 288311
LENGTH: 573
TYPE: DNA
ORGANISM: Human
US-10-027-632-288311

Query Match          58.5%; Score 24; DB 15; Length 573;
Best Local Similarity 75.0%; Pred. No. 2.6;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY      2 CAGTTCGTGAGAGCTGCATACACAGGACTGCTTGAGAC 41
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      270 CAGCTACTGGGAGGCTGCATGTTAGGATTGCTTGAGCC 309

RESULT 12
US-09-764-847-1280/c
; Sequence 1280, Application US/09764847
; Patent No. US20020132767A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC009
CURRENT APPLICATION NUMBER: US/09/764,847
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 2003
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1280
LENGTH: 17335
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-847-1280

Query Match          58.5%; Score 24; DB 11; Length 17335;
Best Local Similarity 75.0%; Pred. No. 3.6;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY      2 CAGTTCGTGAGAGCTGCATACACAGGACTGCTTGAGAC 41
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      5266 CAGCTACTGGGAGGCTGCATGTTAGGATTGCTTGAGCC 5227

RESULT 13
US-10-092-154-1280/c
; Sequence 1280, Application US/10092154
; Publication No. US20030054375A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
```

2 CAGTTGCTGAGAAGCCTGACATACCAGGACTGCCTGAGAC 41



GenCore version 5.1.6  
 Copyright (c) 1993 - 2003 Compugen Ltd.  
 nucleic - nucleic search, using sw model  
 on: July 22, 2003, 11:16:52 ; Search time 1307 Seconds  
 (without alignments)  
 508.045 Million cell updates/sec

U95626-C-AT-42723  
 Effect score: 41  
 quence: 1 tcagtgtcgaagcctga.....ataccagactgcctgagac 41

Identity NUC  
 Gapop 10.0 , Gapext 1.0  
 aligned: 16154066 seqs, 8097743376 residues  
 tal number of hits satisfying chosen parameters: 32308132

imum DB seq length: 0  
 imum DB seq length: 2000000000  
 at-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Labase :

EST.\*  
 1: em\_estba.\*  
 2: em\_esthum.\*  
 3: em\_estin.\*  
 4: em\_estnu.\*  
 5: em\_estov.\*  
 6: em\_estpl.\*  
 7: em\_estro.\*  
 8: em\_htc.\*  
 9: gb\_est1.\*  
 10: gb\_est2.\*  
 11: gb\_htc.\*  
 12: gb\_est3.\*  
 13: gb\_est4.\*  
 14: gb\_est5.\*  
 15: em\_estfun.\*  
 16: em\_estom.\*  
 17: gb\_gss.\*  
 18: em\_gss\_hum.\*  
 19: em\_gss\_inv.\*  
 20: em\_gss\_pln.\*  
 21: em\_gss\_vrt.\*  
 22: em\_gss\_fun.\*  
 23: em\_gss\_mam.\*  
 24: em\_gss\_mus.\*  
 25: em\_gss\_other.\*  
 26: em\_gss\_pro.\*  
 27: em\_gss\_rod.\*

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

## SUMMARIES

sult No.	Score	Query Match	Length	ID	Description
1	27.2	66.3	471	14	BQ300489
2	27.2	66.3	538	10	AM653656
3	25.6	62.4	315	12	BF935457
4	25.6	62.4	337	9	AA369692
5	25.6	62.4	384	12	BF990363
6	25.6	62.4	384	12	BF995389

7	25.6	62.4	384	12	BQ002012
8	25.6	62.4	423	12	BF760350
9	25.6	62.4	470	12	BQ002336
10	25.6	62.4	493	14	BQ300457
c 11	25.6	62.4	509	17	AQ898508
c 12	25.6	62.4	512	12	BF858648
c 13	25.6	62.4	841	14	BQ719041
c 14	25.2	61.5	440	9	AI140599
c 15	25.2	61.5	529	12	BF831477
c 16	25.2	61.5	753	10	AW978551
c 17	25	61.0	352	17	B85322
c 18	25	61.0	368	17	AQ132663
c 19	24	58.5	214	10	BB596733
c 20	24	58.5	277	12	BQ007496
c 21	24	58.5	280	9	AA372587
c 22	24	58.5	309	12	BF931001
c 23	24	58.5	322	12	BF926436
c 24	24	58.5	361	14	BQ345670
c 25	24	58.5	374	10	AW386852
c 26	24	58.5	380	13	BI025579
c 27	24	58.5	399	10	AV730530
c 28	24	58.5	412	17	AQ181316
c 29	24	58.5	424	10	AV729997
c 30	24	58.5	457	17	AQ423186
c 31	24	58.5	470	12	BF904944
c 32	24	58.5	480	12	BF997233
c 33	24	58.5	509	12	BE933206
c 34	24	58.5	524	17	AQ784899
c 35	24	58.5	543	14	BQ323349
c 36	24	58.5	560	17	AQ419838
c 37	24	58.5	561	13	BG986513
c 38	24	58.5	571	14	BQ323273
c 39	24	58.5	571	17	AQ504035
c 40	24	58.5	586	9	AL708735
c 41	24	58.5	592	12	BF819216
c 42	24	58.5	602	10	BE144468
c 43	24	58.5	636	17	AQ616424
c 44	24	58.5	648	17	AG147167
c 45	24	58.5	663	17	AG149846

## ALIGNMENTS

RESULT 1  
 BQ300489  
 LOCUS  
 DEFINITION PMO-KT0041-070201-004-b04 KT0041 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BQ300489  
 VERSION BQ300489.1 GI:20816011  
 KEYWORDS EST.  
 SOURCE human.

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 471)

REFERENCE  
 AUTHORS Dias Neto,E., Garcia Correa,R., Venjovski-Almeida,S., Briones,M.R.,  
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,  
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
 ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
 Simpson,A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

TITLE  
 JOURNAL  
 MEDLINE  
 COMMENT  
 Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001



ORIGIN	Query March	Score 25.6;	DB 12;	Length 315;
	Best Local Similarity	62.4%;		
		77.5%;		
	Pred. NO. 24:			

Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
 2 CAGTTGCTGAGAGCTGACATACAGGACTGCTGAGAC 41  
 |||||  
 246 CAGCTACTGGGAAGCTGACATGAGTGGTGGTGGCC 285

SULT 4  
 369692  
 CUS  
 FINITION EST81420 Prostate gland 1 Homo sapiens cDNA 3' end similar to EST  
 containing Alu repeat, mRNA sequence.  
 CESSION  
 RSION  
 YWORDS  
 URCE  
 ORGANISM

AA369692 337 bp mRNA linear EST 21-APR-1997  
 EST81420 Prostate gland 1 Homo sapiens cDNA 3' end similar to EST  
 containing Alu repeat, mRNA sequence.  
 CESSION  
 RSION  
 YWORDS  
 URCE  
 ORGANISM

AA369692.1 GI:2022225  
 EST.  
 human.  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 337)  
 Adams M.D., Soares M.B., Kerlavage A.R., Fields C. and Venter, J.C.  
 Adams M.D., Soares M.B., Kerlavage A.R., Fields C. and Venter, J.C.  
 Rapid cDNA sequencing (expressed sequence tags) from a  
 directionally cloned human infant brain cDNA library  
 Nat. Genet. 4, 373-380 (1993)  
 94004965  
 Other ESTs: EST81421 THC100858  
 Contact: Kerlavage, AR  
 Bioinformatics  
 The Institute for Genomic Research  
 9712 Medical Center Drive, Rockville, MD 20850 USA  
 Tel: 3018699056  
 Fax: 3018699423  
 Email: arkerlav@tigr.org

For clone availability, additional sequence and expression  
 information related to this EST, please check the TIGR Human Gene  
 Index (<http://www.tigr.org/tdb/hgi/hgi.html>)  
 Seq primer: MJ3-21.

FEATURES  
 source  
 1. 337  
 /organism="Homo sapiens"  
 /db\_xref="AFCC (inhost):174463"  
 /db\_xref="taxon:9606"  
 /clone\_lib="Prostate gland I"  
 /sex="male"  
 /dev\_stage="adult, 21 yrs"  
 /note="Organ: prostate; Vector: pBluescript SK-; Site: 1;  
 EcoRI; Site 2: XhoI"  
 76 a 79 c 84 t 1 others

Query Match 62.4%; Score 25.6; DB 9; Length 337;  
 Best Local Similarity 77.5%; Pred. No. 25;  
 Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
 2 CAGTTGCTGAGAGCTGACATACAGGACTGCTGAGAC 41  
 |||||  
 94 CAGCTACTGGGAAGCTGACATGAGTGGTGGTGGCC 133

SULT 5  
 990363  
 CUS  
 FINITION CMO-GN0163-251000-623-a02 GN0163 Homo sapiens cDNA, mRNA sequence.  
 CESSION  
 RSION  
 YWORDS  
 URCE  
 ORGANISM

BF990363 384 bp mRNA linear EST 23-JAN-2001  
 CMO-GN0163-251000-623-a02 GN0163 Homo sapiens cDNA, mRNA sequence.  
 CESSION  
 RSION  
 YWORDS  
 URCE  
 ORGANISM

BF990363.1 GI:12396688  
 EST.  
 human.  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 384)  
 Dias Neto, E., Garcia Corrae, R., Verjovski-Almeida, S., Briones, M.R.,  
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,

TITLE  
 JOURNAL  
 MEDLINE  
 COMMENT

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 20202663  
 Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil

Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM0&st2=CM0-GN0163-251000-623-a02&st3=2000-10-25&t4=1>)  
 Seq primer: puc 18 forward  
 High quality sequence stop: 383.

FEATURES  
 source  
 1. 384  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="GN0163"  
 /dev\_stage="Adult"

/note="Organ: placenta normal; Vector: puc18; Site: 1; SmaI  
 ; Site 2: SmaI; A mini-library was made by cloning  
 products derived from ORESTES PCR (U.S. Letters Patent  
 application No. 196,716 - Ludwig Institute for Cancer  
 Research) profiles into the puc 18 vector. Reverse  
 transcription of tissue mRNA and cDNA amplification were  
 performed under low stringency conditions."  
 88 a 85 c 104 g 107 t

BASE COUNT  
 ORIGIN  
 Query Match 62.4%; Score 25.6; DB 12; Length 384;  
 Best Local Similarity 77.5%; Pred. No. 26;  
 Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 CAGTTGCTGAGAGCTGACATACAGGACTGCTGAGAC 41  
 |||||  
 Db 137 CAGCTGCTGGGAAGCTGACATGAGTGGTGGTGGCC 176

RESULT 6  
 BF995389 384 bp mRNA linear EST 23-JAN-2001  
 LOCUS  
 DEFINITION CMO-GN0163-041100-623-a02 GN0163 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BF995389  
 VERSION BF995389.1 GI:12401712  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 384)  
 Dias Neto, E., Garcia Corrae, R., Verjovski-Almeida, S., Briones, M.R.,  
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,  
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare  
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
 Simpson, A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 20202663  
 Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil

TITLE  
 JOURNAL  
 MEDLINE  
 COMMENT

Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM0&t2=CM0-GN0163-041100-623-a02&t3=2000-11-04&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 383.  
Location/Qualifiers  
1. .384  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="GN0163"  
/dev\_stage="Adult"  
/note="Organ: placenta normal; Vector: puc18; Site 1: SmaI ; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."  
BASE COUNT 88 a 85 c 104 g 107 t  
ORIGIN  
1. .384  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="GN0163"  
/dev\_stage="Adult"  
/note="Organ: placenta normal; Vector: puc18; Site 1: SmaI ; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."  
Query Match 62.4%; Score 25.6; DB 12; Length 384;  
Best Local Similarity 77.5%; Pred. No. 26;  
Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
2 CAGTTGCTGAGAGCCTGACATACGAGGACTGCCTGAGAC 41  
|||||  
137 CAGCTGCTGGAGGCTGACATGCTAGATTGCTTGAGCC 176  
|||||  
RESULT 7  
BG002012 384 bp mRNA linear EST 24-JAN-2001  
CM0-GN0162-151100-623-a02 GN0162 Homo sapiens cDNA, mRNA sequence.  
BG002012  
BG002012.1 GI:12440900  
EST.  
YWORDS  
URCE human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 384)  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM0&t2=CM0-GN0162-151100-623-a02&t3=2000-11-15&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 383.  
Location/Qualifiers  
1. .384  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="GN0163"  
/dev\_stage="Adult"  
/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."  
BASE COUNT 88 a 85 c 104 g 107 t  
ORIGIN  
1. .384  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="GN0163"  
/dev\_stage="Adult"  
/note="Organ: placenta normal; Vector: puc18; Site 1: SmaI ; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."  
Query Match 62.4%; Score 25.6; DB 12; Length 384;  
Best Local Similarity 77.5%; Pred. No. 26;  
Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
2 CAGTTGCTGAGAGCCTGACATACGAGGACTGCCTGAGAC 41  
|||||  
137 CAGCTGCTGGAGGCTGACATGCTAGATTGCTTGAGCC 176  
|||||  
RESULT 8  
BF760350 423 bp mRNA linear EST 12-JAN-2001  
CM1-CT0605-211200-671-b09 CT0605 Homo sapiens cDNA, mRNA sequence.  
BF760350  
BF760350.1 GI:12108250  
EST.  
KEYWORDS  
SOURCE human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 423)  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM1&t2=CM1-CT0605-211200-671-b09&t3=2000-12-21&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 7  
High quality sequence stop: 423.  
Location/Qualifiers  
1. .423  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="CT0605"  
/dev\_stage="Adult"  
/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM0&t2=CM0-GN0163-041100-623-a02&t3=2000-11-04&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 383.  
Location/Qualifiers  
1. .384  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="GN0163"  
/dev\_stage="Adult"  
/note="Organ: placenta normal; Vector: puc18; Site 1: SmaI ; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."  
BASE COUNT 88 a 85 c 104 g 107 t  
ORIGIN  
1. .384  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="GN0163"  
/dev\_stage="Adult"  
/note="Organ: placenta normal; Vector: puc18; Site 1: SmaI ; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."  
Query Match 62.4%; Score 25.6; DB 12; Length 384;  
Best Local Similarity 77.5%; Pred. No. 26;  
Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
2 CAGTTGCTGAGAGCCTGACATACGAGGACTGCCTGAGAC 41  
|||||  
137 CAGCTGCTGGAGGCTGACATGCTAGATTGCTTGAGCC 176  
|||||  
RESULT 8  
BF760350 423 bp mRNA linear EST 12-JAN-2001  
CM1-CT0605-211200-671-b09 CT0605 Homo sapiens cDNA, mRNA sequence.  
BF760350  
BF760350.1 GI:12108250  
EST.  
KEYWORDS  
SOURCE human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 423)  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM1&t2=CM1-CT0605-211200-671-b09&t3=2000-12-21&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 7  
High quality sequence stop: 423.  
Location/Qualifiers  
1. .423  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="CT0605"  
/dev\_stage="Adult"  
/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

```

% COUNT      119 a      92 c      110 g      102 t
GIN

Query Match      62.4%; Score 25.6; DB 12; Length 423;
Best Local Similarity 77.5%; Pred. No. 27;
Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

2 CAGTTGCTGAGAGCTGACATACACAGGACTGCTGAGAC 41
|||||
253 CAGCTACTGGGAGGCTGACATGTAGGACTGCTTGAGCC 292

%ULT 9
%02336
US      470 bp      mRNA      linear      EST 24-JAN-2001
INITIATION QV4-GN0250-161100-543-c01 GN0250 Homo sapiens cDNA, mRNA sequence.
SSION      EG002336
%ION      EG002336.1 GI:12441546
WORDS      EST.
ACE      human.
%ANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 470)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
EDLINE      20202663
MENT      Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV4&t2=QV4-GN0250-
161100-543-c01&t3=2000-11-18&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 2
High quality sequence stop: 470.
High quality sequence stop: 470.
Location/Qualifiers
1..470
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GN0250"
/dev_stage="Adult"
/note="Organ: placenta normal; Vector: puc18; Site 1: SmaI
; Site 2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the pUC 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."
% COUNT      119 a      104 c      125 g      122 t
GIN

Query Match      62.4%; Score 25.6; DB 12; Length 470;
Best Local Similarity 77.5%; Pred. No. 28;
Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

2 CAGTTGCTGAGAGCTGACATACACAGGACTGCTGAGAC 41
|||||
62 CAGCTACTGGGAGGCTGACATGTAGGACTGCTTGAGCC 101

%ULT 9
%02336
US      470 bp      mRNA      linear      EST 24-JAN-2001
INITIATION QV4-GN0250-161100-543-c01 GN0250 Homo sapiens cDNA, mRNA sequence.
SSION      EG002336
%ION      EG002336.1 GI:12441546
WORDS      EST.
ACE      human.
%ANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 470)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
EDLINE      20202663
MENT      Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV4&t2=QV4-GN0250-
161100-543-c01&t3=2000-11-18&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 2
High quality sequence stop: 470.
High quality sequence stop: 470.
Location/Qualifiers
1..470
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GN0250"
/dev_stage="Adult"
/note="Organ: placenta normal; Vector: puc18; Site 1: SmaI
; Site 2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the pUC 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."
BASE COUNT      137 a      126 c      120 g      110 t
ORIGIN

Query Match      62.4%; Score 25.6; DB 14; Length 493;
Best Local Similarity 77.5%; Pred. No. 28;
Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

2 CAGTTGCTGAGAGCTGACATACACAGGACTGCTGAGAC 41
|||||
187 CAGCTACTGGGAGGCTGACATGTAGGACTGCTTGAGCC 226

%ULT 11
%AQ898508/c
LOCUS      AQ898508
DEFINITION HS 3127 B2 C10 T7C CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone Plate=3127 Col=20 Row=F, DNA sequence.
ACCESSION AQ898508
VERSION      AQ898508.1 GI:6354698
KEYWORDS      GSS.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 493)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
EDLINE      20202663
MENT      Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PMO&t2=PMO-KT0041-
030201-003-h02&t3=2001-02-03&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 14
High quality sequence stop: 467.
High quality sequence stop: 467.
Location/Qualifiers
1..493
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="KT0041"
/dev_stage="Adult"
/note="Organ: bladder tumor; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

```

Rukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 509)  
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.

## TITLE

Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

## JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

## EDLINE

99380589

## MENT

Contact: Mahairas GG, Wallace JC, Hood L

High Throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Clones may be purchased from Research Genetics (info@resgen.com).

BAC end Web Server: http://www.htsc.washington.edu

Plate: 3127 row: F column: 20

Seq primer: F7

Class: BAC ends

High quality sequence stop: 509.

## FEATURES

source

1..509

Location/Qualifiers

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="CIT Approved Human Genomic Sperm Library D"

/sex="male"

/note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in E-Coli DH10B"

125 a 121 c 103 g 154 t 6 others

## BASE COUNT

Query Match 62.4%; Score 25.6; DB 17; Length 509;

Best Local Similarity 77.5%; Pred. No. 29;

Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 CAGTTGCTGAGAGCTGACATACAGGACTGCTGAGAC 41

DB 237 CAGTTACTGAGAGGCTGAGGTAGGAGGATGCTTGAGCC 198

RESULT 12

858648

CUS

DEFINITION RC5-FT0194-271100-022-A06 FT0194 Homo sapiens cDNA, mRNA sequence.

CESSION BF858648

ASTON BF858648.1 GI:12246392

WORDS EST.

human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 512)

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare

M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

FEATURES

source

1..841

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="IMAGE:6189820"

/sex="male"

/tissue\_type="sympathetic trunk"

/dev\_stage="adult, 16 yr"

/lab\_host="DH10B"

/note="Vector: pCMV-SPORT6 (Life Technologies); Site 1:

Not1; Site 2: SalI; cDNA made by oligo-dT priming.

Directionally cloned using the following adaptors:

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPSP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC5&l2=RC5-FT0194-

271100-022-A06&l3=2000-11-27&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 73

High quality sequence stop: 506.

Location/Qualifiers

1..512

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="FT0194"

/dev\_stage="Adult"

/note="Organ: prostate tumor; Vector: puc18; Site 1: SmaI;

Site 2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the pUC 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

BASE COUNT 146 a 110 c 138 g 118 t

ORIGIN

Query Match 62.4%; Score 25.6; DB 12; Length 512;

Best Local Similarity 77.5%; Pred. No. 29;

Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 CAGTTGCTGAGAGCTGACATACAGGACTGCTTGAGAC 41

DB 257 CAGCTACTGGAAGGCTGACATGCTAGGACTGCTTGAGCC 296

RESULT 13

BQ719041/c

LOCUS

DEFINITION BQ719041 841 bp mRNA linear EST 16-JUL-2002

AGENCOURT 8100912 Lupski\_sympathetic\_trunk Homo sapiens cDNA clone

IMAGE:6189820 5', mRNA sequence.

BQ719041

ACCESSION BQ719041.1 GI:21857938

VERSION EST.

KEYWORDS

SOURCE

human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 841)

NIH-MGC http://mgi.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs@mail.nih.gov

Tissue Procurement: Dr. James R. Lupski

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LHAM13587 row: k column: 05

High quality sequence stop: 615.

Location/Qualifiers

1..841

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="IMAGE:6189820"

/sex="male"

/tissue\_type="sympathetic trunk"

/dev\_stage="adult, 16 yr"

/lab\_host="DH10B"

/note="Vector: pCMV-SPORT6 (Life Technologies); Site 1:

Not1; Site 2: SalI; cDNA made by oligo-dT priming.

Directionally cloned using the following adaptors:

5'-TCGACCCAGCGCTCCG-3' and  
 5'-GACTAGTCTCAGATCCGAGCGCGCCT(15)-3'. Size selected >  
 1 kb for average insert length 1.9 kb. This is a primary  
 library, non-amplified. Library constructed by Life  
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor  
 College of Medicine); available through Life  
 Technologies."

QUERY COUNT 187 a 221 c 207 g 226 t  
 GIN

Query Match 62.4%; Score 25.6; DB 14; Length 841;  
 Best Local Similarity 77.5%; Pred. No. 34;  
 Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

2 CAGTGTCTGAGAGCTGACATACCCAGGACTCGTCGAGAC 41  
 |||||  
 361 CAGTGTCTGAGAGCTGACATGCTAGGACTGCTGAGCC 322  
 |||||

AL140599 440 bp mRNA linear EST 29-OCT-1998  
 ge05e09.xl Soares\_testis\_NHT Homo sapiens cDNA clone IMAGE:1738120  
 3' similar to contains Alu repetitive element?, mRNA sequence.  
 AL140599  
 AL140599.1 GI:3648056

WORD SOURCE

ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 440)

NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo  
 , Ph.D.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CCGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 www.bio.llnl.gov/bbrp/image/image.html

Insert Length: 377 Std Error: 0.00

Seq primer: -40ml3 fwd. ET from Amersham

High quality sequence stop: 427.

Location/Qualifiers

source

1. 440  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1738120"  
 /clone\_lib="Soares\_testis\_NHT"  
 /sex="male"  
 /lab\_host="DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified  
 polylinker; site\_1: Not I; Site 2: Eco RI; 1st strand cDNA  
 was prepared from mRNA obtained from Clontech Laboratories  
 , Inc., and primed with a Not I - oligo(dT) primer [5'  
 TGTTACCAATCGTAAGTGGAGCGCGCCCAATTTTTTTTTTTT 3'].  
 Double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified pT73 vector. Library  
 went through one round of normalization to Cot5, and was  
 constructed by Bento Soares and M. Fatima Bonaldo."  
 103 a 115 c 60 g 161 t 1 others

Query Match 61.5%; Score 25.2; DB 9; Length 440;  
 Best Local Similarity 78.9%; Pred. No. 38;  
 Matches 30; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 TCAGTGTCTGAGAGCTGACATACCCAGGACTGCTGTA 38  
 |||||  
 DB 101 TCAGTGTCTGAGAGCTGACGTAGGAGATTGCTTGA 64  
 |||||

RESULT 15

LOCUS BF891477

DEFINITION

BF891477

VERSION

BF891477.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 529)

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare

,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

CONTACT: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM3&t2=PM3-MT0110-

241000-010-g07&t3=2000-10-24&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 49

High quality sequence stop: 529.

Location/Qualifiers

1. 529

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="MT0110"

/dev\_stage="Adult"

/note="Organ: marrow; Vector: puc18; Site 1: SmaI; Site 2:

SmaI; A mini-library was made by cloning products derived

from ORESTES PCR (U.S. Letters Patent application No. 196

,716 - Ludwig Institute for Cancer Research) profiles

into the pUC 18 vector. Reverse transcription of tissue

mRNA and cDNA amplification were performed under low

stringency conditions."

BASE COUNT 161 a 102 c 110 g 156 t

ORIGIN

Query Match 61.5%; Score 25.2; DB 12; Length 529;

Best Local Similarity 78.9%; Pred. No. 41;

Matches 30; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 TCAGTGTCTGAGAGCTGACATACCCAGGACTGCTGTA 38

DB 152 TCAGTGTCTGAGAGCTGACGTAGGAGATTGCTTGA 189

Search completed: July 22, 2003, 12:02:13

Job time : 1311 secs



GenCove version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

nucleic - nucleic search, using sw model  
on: July 22, 2003, 11:16:54 ; Search time 985 Seconds  
(without alignments)  
1211.386 Million cell updates/sec

ie: U95626-A-AT-42723  
fect score: 41  
ence: 1 tcagtgtcgaagcctga.....ataccaggactgcctgagac 41

ing table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

ched: 2054640 seqs, 14551402878 residues

l number of hits satisfying chosen parameters: 4109280

'mum DB seq length: 0

'mum DB seq length: 2000000000

u-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

abase :

GenEmbl:\*

1: gb.ba.\*

2: gb.btg.\*

3: gb.in.\*

4: gb.ov.\*

5: gb.pat.\*

6: gb.ph.\*

7: gb.pl.\*

8: gb.pr.\*

9: gb.ro.\*

10: gb.sts.\*

11: gb.sy.\*

12: gb.un.\*

13: gb.vi.\*

14: gb.ba.\*

15: em.fun.\*

16: em.fun.\*

17: em.fun.\*

18: em.in.\*

19: em.mu.\*

20: em.om.\*

21: em.or.\*

22: em.ov.\*

23: em.pat.\*

24: em.ph.\*

25: em.pl.\*

26: em.ro.\*

27: em.sts.\*

28: em.un.\*

29: em.vi.\*

30: em.htg.hum.\*

31: em.htg.inv.\*

32: em.htg.other.\*

33: em.htg\_mus.\*

34: em.htg\_pln.\*

35: em.htg\_rod.\*

36: em.htg\_man.\*

37: em.htg\_vrt.\*

38: em.sy.\*

39: em.htgo.hum.\*

40: em.htgo\_mus.\*

41: em.htgo\_other.\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	39.4	96.1	583	9	HSMCPA01	U80923 Human monoc
2	39.4	96.1	2082	9	AF068265	AF068265 Homo sapi
3	39.4	96.1	143068	6	AX335952	AX335952 Sequence
4	39.4	96.1	143068	9	HSU95626	U95626 Homo sapien
5	39.4	96.1	185437	9	AC098613	AC098613 Homo sapi
6	27.2	66.3	120793	2	AF322450	AF322450 Homo sapi
7	27.2	66.3	133544	9	AC100748	AC100748 Homo sapi
8	27.2	66.3	163035	2	AC006405	AC006405 Homo sapi
9	26.6	64.9	82601	9	AL136971	AL136971 Human DNA
10	26.6	64.9	142441	2	HSU344H20	AL132710 Homo sapi
11	26.6	64.9	202750	2	AC116942	AC116942 Pan trogl
12	25.6	62.4	59155	2	AC104983	AC104983 Homo sapi
13	25.6	62.4	110394	9	AC003086	AC003086 Homo sapi
14	25.6	62.4	111312	9	AC011248	AC011248 Homo sapi
15	25.6	62.4	142082	2	AC093137	AC093137 Pan trogl
16	25.6	62.4	159365	9	AL392111	AL392111 Human DNA
17	25.6	62.4	171267	2	AC040979	AC040979 Homo sapi
18	25.6	62.4	176738	9	AC090960	AC090960 Homo sapi
19	25.6	62.4	178451	9	AL139396	AL139396 Human DNA
20	25.6	62.4	178942	9	AL357520	AL357520 Human DNA
21	25.6	62.4	181755	2	AC013283	AC013283 Homo sapi
22	25.6	62.4	190948	2	AC093138	AC093138 Pan trogl
23	25.6	62.4	208392	2	AC084876	AC084876 Homo sapi
24	25.6	62.4	210736	2	AC092792	AC092792 Homo sapi
25	25.6	62.4	214304	2	AC026112	AC026112 Homo sapi
26	25.2	61.5	132463	2	AC111775	AC111775 Rattus no
27	25.2	61.5	187803	10	AL596252	AL596252 Mouse DNA
28	25	61.0	41615	9	AC005625	AC005625 Homo sapi
29	25	61.0	68116	2	AC103970	AC103970 Homo sapi
30	25	61.0	73235	2	AC103725	AC103725 Homo sapi
31	25	61.0	78986	2	AC115626	AC115626 Homo sapi
32	25	61.0	112204	9	AL136439	AL136439 Human DNA
33	25	61.0	113836	9	AL133174	AL133174 Human DNA
34	25	61.0	116845	9	AC008819	AC008819 Homo sapi
35	25	61.0	117583	9	AC034238	AC034238 Homo sapi
36	25	61.0	120689	9	CNS01RGS	AL159191 Human chr
37	25	61.0	133322	2	AC091845	AC091845 Homo sapi
38	25	61.0	143506	9	AC092833	AC092833 Homo sapi
39	25	61.0	145871	9	AL391601	AL391601 Human DNA
40	25	61.0	151122	2	AF239614	AF239614 Homo sapi
41	25	61.0	152141	2	AC090189	AC090189 Homo sapi
42	25	61.0	155428	9	AC009682	AC009682 Homo sapi
43	25	61.0	159185	2	AC013550	AC013550 Homo sapi
44	25	61.0	160979	2	AC055879	AC055879 Homo sapi
45	25	61.0	164125	9	AC018693	AC018693 Homo sapi

ALIGNMENTS

RESULT 1  
HSMCPA01  
LOCUS Human monocyte chemoattractant protein 1 receptor gene, 5' region.  
DEFINITION U80923  
ACCESSION U80923.1 GI:1773031  
VERSION 1 of 2  
KEYWORDS Homo sapiens.  
SEGMENT Homo sapiens.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 583)  
AUTHORS Wong,L.M., Myers,S.J., Tsou,C.L., Gosling,J., Arai,H. and Charo,I.F.

Pred. No. is the number of results predicted by chance to have a



TITLE Organization and differential expression of the human monocyte chemoattractant protein 1 receptor gene. Evidence for the role of the carboxyl-terminal tail in receptor trafficking

J. Biol. Chem. 272 (2), 1038-1045 (1997)  
 97150864  
 MEDLINE  
 PUBMED  
 8995400  
 2 (bases 1 to 583)  
 MYERS, S.J. and CHARO, I.F.  
 Direct Submission  
 Submitted (04-DEC-1996) Pharmacology, Emory University, 1510 Clifton Road, Atlanta 30322, USA

TURES Location/Qualifiers  
 1..583  
 source

LOCUS 157 a 136 c 156 g 134 t  
 DEFINITION /organism="Homo sapiens"  
 /db\_xref="taxon:9606"

Query Match 96.1%; Score 39.4; DB 9; Length 583;  
 Best Local Similarity 97.6%; Pred. No. 5.2e-05;  
 Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 TCAGTTGCTGAGAGCCTGAAATACACAGGACTGCTGAGAC 41  
 62 TCAGTTGCTGAGAGCCTGACATACACAGGACTGCTGAGAC 102

AF068265 2082 bp DNA linear PRI 20-APR-1999  
 US Homo sapiens monocyte chemoattractant protein 1 receptor (CCR2)  
 gene promoter and mRNA, partial sequence.

AF068265  
 AF068265.1 GI:4587865  
 Homo sapiens.  
 RCE  
 Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Yamamoto, K., Takeshima, H., Hamada, K., Nakao, M., Kino, T., Nishi, T.,  
 Kochi, M., Kuratsu, J., Yoshimura, T. and Ushio, Y.

TITLE Cloning and functional characterization of the 5'-flanking region of the human monocyte chemoattractant protein-1 receptor (CCR2) gene. Essential role of 5'-untranslated region in tissue-specific expression

J. Biol. Chem. 274 (8), 4646-4654 (1999)  
 99143121  
 PUBMED  
 9988701  
 2 (bases 1 to 2082)  
 Yamamoto, K., Takeshima, H., Hamada, K., Nakao, M., Kino, T., Nishi, T.,  
 Kochi, M., Kuratsu, J.-I., Yoshimura, T. and Ushio, Y.

Direct Submission  
 Submitted (26-MAY-1998) Neurosurgery, Kumamoto University School of Medicine, Honjo, 1-1-1, Kumamoto 860-8556, Japan

TURES Location/Qualifiers  
 1..2082  
 source

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="3"  
 /map="3p"

gene 1..>2082  
 /gene="CCR2"

promoter 1..1697  
 /gene="CCR2"

CAAT\_signal 1570..1574  
 /gene="CCR2"

TATA\_signal 1566..1672  
 /gene="CCR2"

mRNA 1698..>2012  
 /gene="CCR2"

5'UTR 1698..2012  
 /product="monocyte chemoattractant protein 1 receptor"

/gene="CCR2"  
 BASE COUNT 650 a 439 c 433 g 560 t  
 ORIGIN

Query Match 96.1%; Score 39.4; DB 9; Length 2082;  
 Best Local Similarity 97.6%; Pred. No. 4.6e-05;  
 Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCAGTTGCTGAGAGCCTGAAATACACAGGACTGCTGAGAC 41  
 DB 1959 TCAGTTGCTGAGAGCCTGACATACACAGGACTGCTGAGAC 1999

RESULT 3  
 AX335952 143068 bp DNA linear PAT 09-JAN-2002  
 DEFINITION Sequence 6461 from Patent WO0194629.  
 ACCESSION AX335952  
 VERSION AX335952.1 GI:18126671  
 KEYWORDS SOURCE

human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1  
 AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,  
 Horrigan, S., Soppet, D.R. and Weaver, Z.

TITLE Cancer gene determination and therapeutic screening using signature gene sets

JOURNAL Patent: WO 0194629-A 6461 13-DEC-2001;  
 Avalon Pharmaceuticals (US)

FEATURES Location/Qualifiers  
 1..143068  
 source /organism="Homo sapiens"  
 /db\_xref="taxon:9606"

BASE COUNT 41194 a 30122 c 32403 g 39349 t  
 ORIGIN

Query Match 96.1%; Score 39.4; DB 6; Length 143068;  
 Best Local Similarity 97.6%; Pred. No. 3e-05;  
 Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCAGTTGCTGAGAGCCTGAAATACACAGGACTGCTGAGAC 41  
 DB 42703 TCAGTTGCTGAGAGCCTGACATACACAGGACTGCTGAGAC 42743

HSU95626 143068 bp DNA linear PRI 16-MAY-1997  
 LOCUS Homo sapiens ccr2b (ccr2), ccr2a (ccr2), ccr5 (ccr5) and ccr6  
 DEFINITION (ccr6) genes, complete cds, and lactoferrin (lactoferrin) gene,  
 partial cds, complete sequence.

U95626  
 U95626.1 GI:2104517  
 KEYWORDS HTG.

SOURCE Homo sapiens.  
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 143068)  
 AUTHORS McCombie, W.R., Wilson, R., Chen, E., Gibbs, R., Zuo, L., Johnson, D.,  
 Nhan, M., Parnell, L., Dedhia, N., Ansari, A., Mardis, E., Schutz, K.,  
 Gnoj, L., de la Bastide, M., Kaplan, N., Greco, T., Touchman, J.,  
 Muzny, D., Chen, C.-N., Evans, C., Fitzgerald, M., See, L.H., Tang, M.,  
 Porcel, B.M., Dragan, Y., Giacalone, J., Pae, A., Powell, E.,  
 Solinsky, K.A., Desilva, U., Diaz-Perez, S., Zhou, X., Yu, Y.,  
 Watanabe, M., Doggett, N., Garcia, D. and Sagripanti, J.-L.

Human BAC clone 110P12  
 Unpublished (1997)

TITLE  
 JOURNAL  
 REFERENCE 2 (bases 1 to 143068)

AUTHORS McCombie, R.W., Wilson, R., Chen, E., Gibbs, R., Zuo, L., Johnson, D.,  
 Nhan, M., Parnell, L., Dedhia, N., Ansari, A., Mardis, E., Schutz, K.,



CAPGSDPRNLCAICIDQENKCPVNSERNYYGTGAFRCLEAGADVAFFVDVTV  
 LQNTDNNDAWAKDLALFALCLDGRKKPVTEARSLHMAPNAHVSRNDKVER  
 LKQVLLHQAKFGKNGSDPKFCFLQSETKNLLFNDETECLARLHGKTYEKLGPQ  
 YVAGITNLKCKSTSPLEACEFLRK"  
 (SE COUNT 41194 a 30122 c 32403 g 39349 t  
 IGIN

Query Match 96.1%; Score 39.4; DB 9; Length 143068;  
 Best Local Similarity 97.6%; Pred. No. 3e-05;  
 Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 TCAGTCTGAGAGCTGAAATACGAGACTGCTGAGAC 41

42703 TCAGTCTGAGAGCTGACTACGAGACTGCTGAGAC 42743

SULT 5  
 CUS AC098613 185437 bp DNA linear PRI 01-AUG-2002  
 FINITION Homo sapiens chromosome 3 clone RP11-24F11, complete sequence.

CESSION AC098613  
 RSION AC098613.2 GI:22038607

YWORDS HTG.

URCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 185437)

Kaul,R.K., Olson,M.V., Raymond,C. and Haugen,E.D.  
 Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,  
 Saenphimmachak,C., Phelps,K.A., Buckley,D., Kibukawa,M., Raymond,C.  
 and Haugen,E.D.

Direct Submission

Unpublished

2 (bases 1 to 185437)

Kaul,R.K., Olson,M.V., Raymond,C. and Haugen,E.D.

Direct Submission

Submitted (26-OCT-2001) Genome Center, University of Washington,  
 Box 352145, Seattle, WA 98195, USA

3 (bases 1 to 185437)

Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,  
 Saenphimmachak,C., Phelps,K.A., Buckley,D., Kibukawa,M., Raymond,C.  
 and Haugen,E.D.

Direct Submission

Submitted (01-AUG-2002) Genome Center, University of Washington,  
 Box 352145, Seattle, WA 98195, USA

On Aug 1, 2002 this sequence version replaced gi:16445164.

----- Genome Center

Center: University of Washington Genome Center

Center Code: UWGC

Web site: <http://www.genome.washington.edu>

Contact: [uwgctgs@u.washington.edu](mailto:uwgctgs@u.washington.edu)

----- Project Information

Center project name: chr-3

Center clone name: RP11-24F11 (bc0137)

----- Summary Statistics

Sequencing vector: plasmid; 100% of reads

Chemistry: Dye-terminator ET; 93% of reads

Chemistry: Dye-terminator Big Dye; 7% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 184860 bases at least Q40

Consensus quality: 185398 bases at least Q30

Consensus quality: 185435 bases at least Q20

Insert size: 185437; sum-of-contigs

Quality coverage: 7.6x in Q20 bases; sum-of-contigs

-----

Overlapping Sequences:

5': BAC-110P12 U95626, 111014-bp overlap

3': RP11-509I21 (UWGC:bc0454) AC104304, 61294-bp overlap

-----

Sequence Quality Assessment:

This entry has been annotated with sequence quality  
 estimates computed by the Phrap assembly program.

All manually edited bases have been reduced to quality zero.

Quality levels above 40 are expected to have less than  
 1 error in 10,000 bp.  
 Base-by-base quality values are not generally visible from the  
 Genbank flat file format but are available as part  
 of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted:  
 all regions were either double-stranded or sequenced with an  
 alternate chemistry or covered by high quality data (i.e., Phred  
 quality >= 30); an attempt was made to resolve all sequencing  
 problems, such as compressions and repeats; all regions were  
 covered by at least one plasmid subclone or more than one M13  
 subclone; and the assembly was confirmed by restriction digest.

#### Sequence Validation:

This sequence has been validated by Multiple Complete Digest  
 fingerprinting. Comparison of the experimentally derived digest  
 fragments with sequence-predicted fragments is given below.  
 The electronically-digested sequence consists of both insert and  
 vector, in order to accurately represent the entire circular BAC.  
 Small fragments below a variable cutoff (approximately 400-800 bp)  
 are not resolved in the fingerprint and hence do not appear  
 in the table. There are no significant remaining discrepancies  
 between the experimental and predicted values. Uniquely ordered  
 fragments are separated by dashed lines.

SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt
-----	-----	-----	-----	-----	-----
11786	12040	8696	8837	9531	9423
-----	-----	-----	-----	-----	-----
2067	2065	6	<800	6382	6501
-----	-----	-----	-----	-----	-----
5681	5720	1846	1824	512	<800
-----	-----	-----	-----	-----	-----
3716	3953	4052	3986	449	<800
-----	-----	-----	-----	-----	-----
875	897	1159	1130	6692	6501
-----	-----	-----	-----	-----	-----
910	897	54	<800	10449	10229
-----	-----	-----	-----	-----	-----
2169	2215	560	<800	953	1025
-----	-----	-----	-----	-----	-----
1705	1653	2287	2267	3124	3176
-----	-----	-----	-----	-----	-----
5763	5720	4905	4891	1054	1025
-----	-----	-----	-----	-----	-----
5844	5720	3049	3078	3985	3998
-----	-----	-----	-----	-----	-----
5061	5001	1377	1394	1100	1025
-----	-----	-----	-----	-----	-----
2625	2640	9903	9772	124	<800
-----	-----	-----	-----	-----	-----
725	<800	1022	1021	2671	2693
-----	-----	-----	-----	-----	-----
1173	1161	12606	12503	1948	1968
-----	-----	-----	-----	-----	-----
5747	5720	866	868	83	<800
-----	-----	-----	-----	-----	-----
9995	9684	9817	9772	7455	7474
-----	-----	-----	-----	-----	-----
999	999	3598	3601	1305	1267
-----	-----	-----	-----	-----	-----
4541	4503	452	<800	1047	1025
-----	-----	-----	-----	-----	-----
2950	2981	7549	7635	3299	3301
-----	-----	-----	-----	-----	-----
406	<800	2063	2075	5279	5237
-----	-----	-----	-----	-----	-----
3375	3602	5837	5858	6815	6926
-----	-----	-----	-----	-----	-----

Query Match	96.1%;	Score 39.4;	DB 9;	Length 185437;
Best Local Similarity	97.6%;	Pred. No. 2.9e-05;		
Matches 40;	Conservative	0;	Mismatches 1;	Indels 0;
Gaps	0;			
QY	1	TCAGTTGCTGAGAGCCTGAATACACGAGCTGCTGAGAC	41	
Db	10653	TCAGTTGCTGAGAGCCTGACATACACGAGCTGCTGAGAC	10693	
RESULT 6				
AF322450				
LOCUS	AF322450	120793 bp	DNA	linear
DEFINITION	Homo sapiens chromosome 17 clone BAC407121 map 17p13.3, ***			
SEQUENCING IN PROGRESS ***;	28 unordered pieces.			
ACCESSION	AF322450			
VERSION	AF322450.1	GI:11559855		
KEYWORDS	HTG; HTGS PHASE1.			
SOURCE	Homo sapiens.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	Zhao,X.T., He,M., Wan,D.F., Ye,Y., Qin,W.X., Huang,Y., Zuo,L. and Gu,J.R.			
TITLE	Gene clone on human chromosome 17p13.3			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 120793)			
AUTHORS	Zhao,X.T., He,M., Wan,D.F., Ye,Y., Qin,W.X., Huang,Y., Zuo,L. and Gu,J.R.			
TITLE	Direct Submission			
JOURNAL	Submitted (21-NOV-2000) National Laboratory For Oncogenes & Related Genes, Shanghai Cancer Institute, 25/Ln 2200 Xie-Tu Road, Shanghai 200332, P. R. China			
COMMENT	* NOTE: This is a 'working draft' sequence. It currently consists of 28 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.			
1	608:	contig of 608 bp in length		
*		gap of unknown length		
*	609	1282:	contig of 674 bp in length	
*		gap of unknown length		
*	1283	1733:	contig of 451 bp in length	
*		gap of unknown length		
*	1734	2435:	contig of 702 bp in length	
*		gap of unknown length		
*	2436	2895:	contig of 460 bp in length	
*		gap of unknown length		
*	2896	3643:	contig of 748 bp in length	
*		gap of unknown length		
*	3644	4149:	contig of 506 bp in length	
*		gap of unknown length		
*	4150	4908:	contig of 759 bp in length	
*		gap of unknown length		
*	4909	5741:	contig of 833 bp in length	
*		gap of unknown length		
*	5742	6777:	contig of 1036 bp in length	
*		gap of unknown length		
*	6778	7766:	contig of 989 bp in length	
*		gap of unknown length.		
*	7767	9183:	contig of 1417 bp in length	
*		gap of unknown length		
*	9184	9931:	contig of 748 bp in length	
*		gap of unknown length		
*	9932	12274:	contig of 2343 bp in length	
*		gap of unknown length		
*	12275	16059:	contig of 3785 bp in length	
*		gap of unknown length		
*	16060	19969:	contig of 3910 bp in length	
*		gap of unknown length		

TURES

source

Location/Qualifiers

1. .185437

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="3"

/clone="RP11-24F11"

TURES  
source  
Location/Qualifiers  
1. .185437  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="3"  
/clone="RP11-24F11"

```

* 19970 22905: contig of 2936 bp in length
*      gap of unknown length
* 22906 28210: contig of 5305 bp in length
*      gap of unknown length
* 28211 34020: contig of 5810 bp in length
*      gap of unknown length
* 34021 41704: contig of 7684 bp in length
*      gap of unknown length
* 41705 48302: contig of 6598 bp in length
*      gap of unknown length
* 48303 56789: contig of 8487 bp in length
*      gap of unknown length
* 56790 65996: contig of 9207 bp in length
*      gap of unknown length
* 65997 74434: contig of 8438 bp in length
*      gap of unknown length
* 74435 83698: contig of 9264 bp in length
*      gap of unknown length
* 83699 95956: contig of 12258 bp in length
*      gap of unknown length
* 95957 107953: contig of 11997 bp in length
*      gap of unknown length
* 107954 120793: contig of 12840 bp in length.

```

# FEATURES

```

Location/Qualifiers
1. .120793
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17p13.3"
/clone="EAC407I21"

```

```

SE COUNT 28674 a 31051 c 31727 g 29285 t 56 others
IGIN

```

```

Query Match 66.3%; Score 27.2; DB 2; Length 120793;

```

```

Best Local Similarity 80.0%; Pred. No. 2.3; Indels 0; Gaps 0;
Matches 32; Conservative 0; Mismatches 8;

```

```

2 CAGTGTGTCGAGAGCCTGAAATACAGAGCTGCTGAGAC 41
|||||

```

```

49 CAGTGTGTCGAGAGCCTGAGATGAGAGGATGCTTGAGCC 88
|||||

```

## SULT 7

```

100748
CUS
FINITION Homo sapiens chromosome 17, clone CTD-2231E3, complete sequence.
CESSION AC100748
SION AC100748.2 GI:21306865
WORDS
HTG.

```

## ORGANISM

```

Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

1 (bases 1 to 133544)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.

```

```

Homo sapiens chromosome 17, clone CTD-2231E3
Unpublished

```

```

2 (bases 1 to 133544)

```

```

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,X., Colangelo,M., Collins,S., Collymore,A., Cooke,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Jones,B., Hearford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRoque,K.,
Lamazares,R., Lander,E., Lehoczyk,J., Levine,R., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Melrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Olivier,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,

```

```

Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

```

## TITLE

```

JOURNAL
Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

```

## REFERENCE

```

3 (bases 1 to 133544)

```

## AUTHORS

```

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,X., Colangelo,M., Collins,S., Collymore,A.,
Cooke,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRoque,K.,
Lamazares,R., Lander,E., Lehoczyk,J., Levine,R., Lindblad-Toh,K.,
Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N.,
Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Melrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Olivier,J., Peterson,K., Phunkhang,P., Pierre,N.,
Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C.,
Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,
Schuback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

```

## TITLE

```

JOURNAL
Submitted (01-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

```

## COMMENT

```

On Jun 1, 2002 this sequence version replaced gi:17048115.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

```

```

----- Genome Center

```

```

Center: Whitehead Institute/ MIT Center for Genome Research

```

```

Center code: WIBR

```

```

Web site: http://www-seq.wi.mit.edu

```

```

Contact: sequence_submissions@genome.wi.mit.edu

```

```

----- Project Information

```

```

Center project name: L17901

```

```

Center clone name: 2231_E_3

```

## FEATURES

```

source

```

```

1. .133544
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="CTD-2231E3"
/clone_lib="CITDI Human BAC"
2. .92

```

```

repeat_region

```

```

/rpt_family="AluY"

```

```

repeat_region

```

```

642_705
/rpt_family="L2"

```

```

repeat_region

```

```

954_1030
/rpt_family="AluSg/x"

```

```

repeat_region

```

```

1033_1159
/rpt_family="AluY"

```

```

repeat_region

```

```

1162_1470
/rpt_family="AluY"

```

```

repeat_region

```

```

1474_1578
/rpt_family="L2"

```

```

repeat_region

```

```

1811_2124
/rpt_family="AluSg"

```

```

repeat_region

```

```

2227_2356
/rpt_family="FLAM_C"

```



http://ftp.genome.washington.edu/RM/RepeatMasker.html.  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 12 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

```

1      8815: contig of 8815 bp in length
      gap of unknown length
*      8816 12285: contig of 3470 bp in length
*      12286 40956: contig of 28671 bp in length
*      40957 64393: contig of 23437 bp in length
*      64394 78975: contig of 14582 bp in length
*      78976 96637: contig of 17662 bp in length
*      96638 105967: contig of 9330 bp in length
*      105968 145832: contig of 39865 bp in length
*      145833 147041: contig of 1209 bp in length
*      147042 150498: contig of 3457 bp in length
*      150499 160979: contig of 10481 bp in length
*      160980 163035: contig of 2055 bp in length.
  
```

## FEATURES

## source

```

1. .163035
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /chromosome="17"
   /map="17"
   /clone_lib="RPCI-11 human BAC library"
   /clone="RP3-322A2"
38384 a 43845 c 42661 g 37996 t 149 others
  
```

## SEQUENCE

## IGIN

```

Query Match      66.3%; Score 27.2; DB 2; Length 163035;
Best Local Similarity 80.0%; Pred. No. 2.2;
Matches 32; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
  
```

```

2 CAGTGTCTGAGAGCTGAAATACGAGACTGCTGAGAC 41
|||||
97969 CAGTGTCTGAGAGCTGAGATGAGAGATTGCTGAGGC 98008
  
```

## SULT 9

## CUS

```

AL136971      82601 bp DNA linear PRI 01-AUG-2000
FINITION Human DNA sequence from clone RP3-322A2 on chromosome 6 Contains
GSSs, complete sequence.
  
```

## CESSION

## RSION

## WORDS

## ORCE

## ORGANISM

```

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 82601)
Harley, J.
  
```

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

```

Submitted (26-JUL-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Jun 22, 2000 this sequence version replaced gi:8670574.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
  
```

together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr6  
 RP3-322A2 is from the library RPCI-3 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see http://bacpac.med.buffalo.edu/

VECTOR: pCYPAC2  
 This sequence is the entire insert of clone RP3-322A2 The true right end of clone RP11-11D8 is at 72842 in this sequence.

## FEATURES

## Location/Qualifiers

## source

```

1. .82601
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /chromosome="6"
   /clone="RP3-322A2"
   /clone_lib="RPCI-3"
436. .509
   /note="2 copies 37 mer 98% conserved"
  
```

```

935. .1106
   /note="L1MB3 repeat: matches 5843. .6016 of consensus"
1072. .1155
   /note="L1MB3 repeat: matches 7654. .7740 of consensus"
2213. .2444
   /note="AluYb repeat: matches 75. .302 of consensus"
2592. .2948
   /note="MLTII repeat: matches 62. .396 of consensus"
3071. .3361
   /note="AluSx repeat: matches 1. .291 of consensus"
3362. .3429
   /note="34 copies 2 mer ta 73% conserved"
4115. .4415
   /note="AluSx repeat: matches 1. .303 of consensus"
4468. .5043
   /note="MER34 repeat: matches 6. .545 of consensus"
5062. .5239
   /note="FAM repeat: matches 1. .172 of consensus"
5302. .5424
   /note="MER58B repeat: matches 204. .326 of consensus"
5576. .5871
   /note="AluSx repeat: matches 1. .294 of consensus"
5977. .6039
   /note="MERAD repeat: matches 920. .981 of consensus"
6040. .6350
   /note="AluSx repeat: matches 1. .312 of consensus"
6351. .6409
   /note="MERAD repeat: matches 863. .920 of consensus"
6410. .6714
   /note="AluSx repeat: matches 1. .308 of consensus"
6715. .6760
   /note="MERAD repeat: matches 819. .863 of consensus"
6761. .7076
   /note="AluYb repeat: matches 1. .303 of consensus"
7077. .7522
   /note="MERAD repeat: matches 359. .819 of consensus"
7522. .7750
   /note="MERAD repeat: matches 3. .234 of consensus"
7757. .7868
  
```

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

## repeat\_region

```

/note="MER5A repeat: matches 25. .161 of consensus"
7893. .8107
repeat_region
/note="AluJb repeat: matches 85. .293 of consensus"
9105. .9398
repeat_region
/note="AluSq repeat: matches 1. .294 of consensus"
9476. .9679
repeat_region
/note="TIGGR1 repeat: matches 26. .235 of consensus"
9680. .9856
repeat_region
/note="AluSx repeat: matches 134. .310 of consensus"
9857. .10162
repeat_region
/note="AluSc repeat: matches 1. .307 of consensus"
10163. .10296
repeat_region
/note="AluSx repeat: matches 1. .134 of consensus"
10297. .10489
repeat_region
/note="TIGGR1 repeat: matches 235. .400 of consensus"
10490. .10801
repeat_region
/note="AluSq repeat: matches 1. .311 of consensus"
10802. .11019
repeat_region
/note="TIGGR1 repeat: matches 400. .623 of consensus"
11020. .11117
repeat_region
/note="AluJb repeat: matches 1. .100 of consensus"
11118. .11483
repeat_region
/note="TIGGR1 repeat: matches 1. .371 of consensus"
11484. .11681
repeat_region
/note="AluJb repeat: matches 100. .290 of consensus"
11691. .12100
repeat_region
/note="TIGGR1 repeat: matches 614. .1027 of consensus"
12101. .12405
repeat_region
/note="AluSp repeat: matches 1. .307 of consensus"
12406. .12571
repeat_region
/note="TIGGR1 repeat: matches 1027. .1193 of consensus"
12572. .12871
repeat_region
/note="AluJb repeat: matches 1. .305 of consensus"
12872. .13167
repeat_region
/note="TIGGR1 repeat: matches 1193. .1491 of consensus"
13168. .13474
repeat_region
/note="AluY repeat: matches 1. .307 of consensus"
13475. .13845
repeat_region
/note="TIGGR1 repeat: matches 1491. .1845 of consensus"
13848. .14172
repeat_region
/note="MER65A repeat: matches 1. .334 of consensus"
14272. .14473
repeat_region
/note="TIGGR1 repeat: matches 1836. .2037 of consensus"
14474. .14750
repeat_region
/note="AluSx repeat: matches 1. .303 of consensus"
14751. .14930
repeat_region
/note="TIGGR1 repeat: matches 2037. .2216 of consensus"
14931. .15220
repeat_region
/note="AluJo repeat: matches 1. .307 of consensus"
15221. .15420
repeat_region
/note="TIGGR1 repeat: matches 2216. .2417 of consensus"
15445. .15747
repeat_region
/note="AluJb repeat: matches 1. .302 of consensus"
15525. .15880
misc_feature
/note="match: GSS: Em:AQ132663"
15527. .15878
misc_feature
/note="match: GSS: Em:B85322"
16334. .16607
repeat_region
/note="AluJo repeat: matches 1. .291 of consensus"
17456. .17567
repeat_region
/note="LIME2 repeat: matches 5976. .6083 of consensus"
18008. .18173
repeat_region
/note="AluSx repeat: matches 119. .288 of consensus"
18210. .18520
repeat_region
/note="AluY repeat: matches 2. .308 of consensus"
19267. .19560
repeat_region
/note="AluSx repeat: matches 3. .293 of consensus"
19605. .19898
repeat_region
/note="MSTA repeat: matches 120. .426 of consensus"
19899. .20283
repeat_region
/note="MSTA repeat: matches 1. .388 of consensus"
20284. .20558
repeat_region
/note="AluSx repeat: matches 36. .312 of consensus"
20559. .20598
repeat_region
/note="MSTA repeat: matches 387. .426 of consensus"
20599. .20712
repeat_region
/note="MSTA repeat: matches 10. .120 of consensus"
20730. .21075
repeat_region
/note="MLT1B repeat: matches 1. .386 of consensus"
21251. .21775
repeat_region
/note="MLT1F repeat: matches 1. .535 of consensus"
22248. .22614
repeat_region
/note="MLT1A1 repeat: matches 1. .365 of consensus"
22714. .23256
repeat_region
/note="MLT2B repeat: matches 1. .440 of consensus"
23500. .23796
repeat_region
/note="AluSp repeat: matches 1. .294 of consensus"
23988. .24221
repeat_region
/note="AluJo repeat: matches 1. .231 of consensus"
complement(24774. .25308)
misc_feature
/note="match: GSS: Em:AQ606839"
complement(24826. .25317)
misc_feature
/note="match: GSS: Em:AQ402803"
24851. .25140
repeat_region
/note="AluSx repeat: matches 15. .306 of consensus"
complement(24853. .25302)
misc_feature
/note="match: GSS: Em:AQ756978"
25326. .25632
repeat_region
/note="AluSx repeat: matches 1. .294 of consensus"
complement(25817. .26235)
misc_feature
/note="match: GSS: Em:AQ058462"
26369. .26654
repeat_region
/note="AluSx repeat: matches 1. .286 of consensus"
27278. .27586
repeat_region
/note="AluJo repeat: matches 1. .309 of consensus"
28001. .28269
repeat_region
/note="AluSx repeat: matches 1. .267 of consensus"
complement(28818. .29223)
misc_feature

Query Match 64.9%; Score 26.6; DB 9; Length 82601;
Best Local Similarity 78.0%; Pred. No. 4.2;
Matches 32; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 TCAGTGTCTGAGAGCTGTAATACCAAGGACTGCTTGAGAC 41
|||||
Db 15611 TCAGTGTCTGAGAGCTGTAATACCAAGGACTGCTTGAGAC 15651

RESULT 10
HSJ344H20/c 142441 bp DNA linear HTG 10-JUL-2001
LOCUS Homo sapiens chromosome 6 clone RP3-344H20 map g15-16.1, ***
DEFINITION SEQUENCING IN PROGRESS ***, 8 unordered pieces.
ACCESSION AL132710
VERSION AL132710.16 GI:9931123
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Brown, J.
TITLE Direct Submission
JOURNAL Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
COMMENT On Aug 27, 2000 this sequence version replaced gi:9864412.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: dJ344H20
----- Summary Statistics
Assembly program: XGAP4; version 4.5

```





```

/note="assembly_fragment"
misc_feature 2260..38575
/note="assembly_fragment"
misc_feature 36676..91546
/note="assembly_fragment"
/note="assembly_fragment"
clone_end:SP6
vector_side:right
91647..202750
/note="assembly_fragment"
clone_end:T7
vector_side:left
3E COUNT 51528 a 47600 c 46613 g 56708 t 301 others
IGIN

Query Match 64.9%; Score 26.6; DB 2; Length 202750;
Best local Similarity 78.0%; Pred. No. 3.8;
Matches 32; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

1 TCAGTCTCAGAGGCTGAAATACGAGCTGCTGAGAC 41
|||||
24734 TCAGTCTCAGAGGCTGAAATACGAGCTGCTGAGAC 24774

SULT 12
104983/c
TUS AC104983 59155 bp DNA linear HTG 22-DEC-2001
Homo sapiens chromosome 18 clone RP11-352C3 map 18, LOW-PASS
SEQUENCE SAMPLING.
SESSION AC104983
RSION AC104983.1 GI:117977361
WORDS HTG; HTGS PHASR0.
JRCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
Biren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 18, clone RP11-352C3
Unpublished
2 (bases 1 to 59155)
Biren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Bouckghalter,B.,
Brown,A., Camarata,J., Campoliano,A., Chang,J., Chazaro,B.,
Choquel,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A.,
Cooke,P., DeAvellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Hearford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRoque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McDermid,P., McKernan,K., McPheeters,R., Meldrum,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
TITLE Submitted (22-DEC-2001) Whitehead Institute/MIT Center for Genome
JOURNAL Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

```

----- Project Information  
Center project name: L22883  
Center clone name: 352\_C3  
-----

\* NOTE: This record contains 70 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1 724: contig of 724 bp in length  
2 725 824: gap of 100 bp  
3 825 1582: contig of 758 bp in length  
4 1583 1682: gap of 100 bp  
5 1683 2409: contig of 727 bp in length  
6 2410 2509: gap of 100 bp  
7 2510 3235: contig of 726 bp in length  
8 3236 3335: gap of 100 bp  
9 3336 4090: contig of 755 bp in length  
10 4091 4190: gap of 100 bp  
11 4191 4955: contig of 765 bp in length  
12 4956 5055: gap of 100 bp  
13 5056 5780: contig of 725 bp in length  
14 5781 5880: gap of 100 bp  
15 5881 6643: contig of 763 bp in length  
16 6644 6743: gap of 100 bp  
17 6744 7496: contig of 753 bp in length  
18 7497 7596: gap of 100 bp  
19 7597 8348: contig of 752 bp in length  
20 8349 8448: gap of 100 bp  
21 8449 9217: contig of 769 bp in length  
22 9218 9317: gap of 100 bp  
23 9318 10093: contig of 776 bp in length  
24 10094 10193: gap of 100 bp  
25 10194 10924: contig of 731 bp in length  
26 10925 11024: gap of 100 bp  
27 11025 11793: contig of 769 bp in length  
28 11794 11893: gap of 100 bp  
29 11894 12661: contig of 768 bp in length  
30 12662 12761: gap of 100 bp  
31 12762 13525: contig of 764 bp in length  
32 13526 13625: gap of 100 bp  
33 13626 14347: contig of 722 bp in length  
34 14348 14447: gap of 100 bp  
35 14448 15172: contig of 725 bp in length  
36 15173 15272: gap of 100 bp  
37 15273 16007: contig of 735 bp in length  
38 16008 16107: gap of 100 bp  
39 16108 16855: contig of 748 bp in length  
40 16856 16955: gap of 100 bp  
41 16956 17711: contig of 756 bp in length  
42 17712 17811: gap of 100 bp  
43 17812 18568: contig of 757 bp in length  
44 18569 18668: gap of 100 bp  
45 18669 19421: contig of 753 bp in length  
46 19422 19521: gap of 100 bp  
47 19522 20229: contig of 708 bp in length  
48 20230 20329: gap of 100 bp  
49 20330 21061: contig of 732 bp in length  
50 21062 21161: gap of 100 bp  
51 21162 21909: contig of 748 bp in length  
52 21910 22009: gap of 100 bp  
53 22010 22773: contig of 764 bp in length  
54 22774 22873: gap of 100 bp  
55 22874 23623: contig of 750 bp in length  
56 23624 23723: gap of 100 bp  
57 23724 24489: contig of 766 bp in length  
58 24490 24589: gap of 100 bp





```

Query Match      62.4%; Score 25.6; DB 9; Length 110394;
Best Local Similarity 77.5%; Pred. No. 10;
Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

2 CAGTGTCTGAGACCTGGAATACAGACGCTGCTTGAGAC 41
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
34215 CAGCTGCTCAGACGCTGAGTGGAGGAGCTGCTTGAGCC 34254

SULT 14
011248
CDS
FINITION
AC011248.8 GI:7658299 linear PRI 28-APR-2000
Homo sapiens 12 BAC RFLP-14521 (Roswell Park Cancer Institute Human
BAC Library) complete sequence.
AC011248
HTG.
Homo sapiens.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 111312)
Muzny, D.M., Adams, C., Bailey, M., Barbata, J., Blankenburg, K.,
Bodota, B., Bouck, J., Bowie, S., Brooks, A., Buhay, C., Bunac, C.,
Burkett, C., Burrows, J., Carter, M., Chacko, J., Chen, Z., Cox, C.,
David, R., Delgado, O., Deshazo, D., Ding, Y., Domah-Rashid, N.,
Dugan-Rocha, S., Durbin, K.J., Fernandez, C., Ferraguto, D.,
Forcum-Tansey, J., Frantz, P., Ganesh, R., Gorrell, J.H., Gorrell, L.L.,
Guevara, W., Harris, K., Hernandez, J., Hodgson, A., Hogues, M.,
Holloway, C., Hosak, H., Issat, A., Jackson, L.E., Jackson, L., Jia, Y.,
Jones, M., Kelly, S., Kneitz, S., Kondeljewski, N., Kong, Y., Kovar, C.,
Lau, S., Leal, B., Lee, E., Li, Z., Lichtarge, O., Liu, J., Liu, W.,
Logan, O., Lorado, R.J., Lu, J., Lucier, R.J., Marondel, I., Martin, R.,
Martinez, C., McLeod, M.P., Mei, G., Merscher, S., Miller, A.,
Montgomery, K.T., Morgan, M., Morris, S., Nash, S., Nelson, A.,
Nguyen, B., Nguyen, N., Nguyen, S., Osval, G., Parish, B., Paxton, S.,
Payton, R., Perez, L., Pu, L.L., Quiles, M., Reiter, D., Rives, M.,
Samuel, S., Say, J., Scherer, S., Shah, E., Shen, H., Shim, C., Simon, M.,
Sparks, A., Stamps, A., Sugang, R., Taboz, P., Taylor, T., Vasquez, L.,
Vinson, R., Vo, Q., Wahbah, M., Watlington, S., Weinstein, G.,
Weinstock, I.R., Williamson, A., Worley, K., Wren, J., Wrensford, G.,
Xiang, A.M., Yang, R., Yu, W., Zhou, X., Kucherlapati, R., Nelson, D., and
Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 111312)
Worley, K.C.
Direct Submission
Submitted (05-OCT-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 111312)
Worley, K.C.
Direct Submission
Submitted (28-APR-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Apr 28, 2000 this sequence version replaced gi:6996901.
INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green,

```

unpublished.) for Human and Mouse sequences.

Genes and region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

#### QUALSTAT-REPORT-----

```

----- Summary Statistics -----
Contig length: 150286
Phrap values in estimate: 149476
Average error rate (BCM-Phrap estimate): 0.000180006
Fraction of Phrap values less than 40 : 0.0291418
Number of consensus changing edits: 15
Number of N's in consensus : 1

```

```

----- Consensus changing edits -----
Position Original+Context Edited+Context
22441 aaagagaaga (n) gagggctcca aaagagaaga (a) gagggctcca
79231 agtaaggcca (n) amttantn agtaaggcca (t) atttttaag
79233 taaggccana (n) ttantncc taaggccat (t) tttaatgcc
79234 aaggccanan (n) ttantncc aaggccat (t) tttaatgcc
79239 nanmttta (n) tcnctngtga catatttta (a) tgcctcgtga
79241 nanmttta (n) cctngtga tatttta (g) cctcgtgaaa
79245 ttantncc (n) gtgaaaataa tttaatgct (c) gtgaaaataa
79279 agactgtga (n) gctagaatg agactgtga (a) gctagaatg
79440 agctggcca (n) catcgtgaaa agctggcca (a) catcgtgaaa
89520 tctcaaaaa (n) caaaagag tctcaaaaa (a) caaaagag
94117 gtttccag (n) tctggcaca gtttccag (t) tctggcaca
95790 ttaagttcc (n) tctcaaaaa ttaagttcc (a) tctcaaaaa
111017 ttaagttcc (n) tcaagtgaa ttaagttcc (a) tcaagtgaa
115887 atgcatata (n) caacaacaa atgcatata (a) caacaacaa
129846 ggaggctgaa (n) cgagtggatc ggaggctgaa (g) cgagtggatc

```

```

----- Distribution of Quality < 40 Bases -----
1000| * *
900| * *
800| * *
700| * *
600| * *
500| * *
400| * *
300| * *
200| * *
100| * *
0| * *
# bases 5 10 15 20 25 30 35 40
Phrap Value Range

```

Version: 1.01 qxf0.

```

repeat_region      6348..6648
/rpt_family="AluSg"
complement(6649..7184)
/rpt_family="MER4A"
8486..8703
/rpt_family="L2"
8729..9122
/rpt_family="L2"
9213..9497
/rpt_family="L2"
9498..9798
/rpt_family="AluY"
9799..9887
/rpt_family="L2"
complement(10707..12039)
/rpt_family="L1MA4"
12047..12155
/rpt_family="L2"
complement(12165..12729)
/rpt_family="L2"
12702..12847
/rpt_family="L2"
1332..13486
/standard_name="Cdalfh02"
/db_xref="GDB:446262"
repeat_region      13505..13528

Query Match      62.4%; Score 25.6; DB 9; Length 111312;
Best Local Similarity 77.5%; Pred. NO. 10;
Matches 31; Conservative 0; Mismatches 9; Indels 0;

2 CAGTTCGTGAGAGCGTGAATAACACGAGGACTGCCTTGAGAC 41
70300 CAGTTCATGTGAGAGCGTCAACTACGAGGACTGCTTGAGAC 70339

```

```

----- Summary Statistics -----
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.950319
Consensus quality: 141565 bases at least Q40
Consensus quality: 141698 bases at least Q30
Consensus quality: 141749 bases at least Q20
Insert size: 113000; agarose-fp
Insert size: 100000; pulse-field-gel
Insert size: 141782; sum-of-contigs
Quality coverage: 13.44x in Q20 bases; agarose-fp
Quality coverage: 15.19x in Q20 bases; pulse-field-gel
Quality coverage: 10.71x in Q20 bases; sum-of-contigs
-----

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1 47102: contig of 47102 bp in length
* 47103 47202: gap of unknown length
* 47203 86175: contig of 38973 bp in length
* 86176 86276: gap of unknown length
* 86276 118675: contig of 32400 bp in length
* 118676 118775: gap of unknown length
* 118776 142082: contig of 23307 bp in length.
*
* Location/Qualifiers
*   1..142082
*   /organism="Pan troglodytes"
*   /db_xref="taxon:9598"
*   /clone="RP43-144H14"
*   /clone_lib="RP43"
*   1..47102
*   /note="assembly_fragment"
*   clone_end:T7
*   vector_side:left"
*   12236..142082
*   /note="clone overlaps with GenBank Accession Number
*   AC093138 clone RP43-154A20 (center project name ann)"
*   47203..86175
*   /note="assembly_fragment"
*   86276..118675
*   /note="assembly_fragment"
*   118776..142082
*   /note="assembly_fragment"
*   clone_end:SP6
*   vector_side:right"
*   132713..142082
*   /note="clone overlaps with GenBank Accession Number
*   AC093189 clone RP43-17P3 (center project name ann)"
*
* SE COUNT 42473 a 27490 c 27247 g 44571 t 301 others
* IGIN
*
* Query Match 62.4%; Score 25.6; DB 2; Length 142082;
* Best Local Similarity 77.5%; Pred. No. 9.9;
* Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
*
* 2 CAGTTCCTGAGAGCCTGAATACAGGACTGCTGAGAC 41
* ||||| ||||| ||||| ||||| ||||| |||||
* 38952 CAGCTGCTCAGAGACTGAGTAGGAGGACTGCTTGAGCC 38991

```

arch completed: July 22, 2003, 13:43:36  
 b time : 989 secs

GenCore version 5.1.6  
 Copyright (C) 1993 - 2003 CompuGen Ltd.  
 nucleic - nucleic search, using sw model  
 on: July 22, 2003, 11:15:53 ; Search time 170.5 Seconds  
 (without alignments)  
 541.536 Million cell updates/sec  
 file: U95626-A-AT-42723  
 effect score: 41  
 sequence: 1 tcagtgcctgagacgctga.....ataccaggactgcctgagac 41  
 oring table: IDENTITY\_NUC  
 Gapop 10.0 , Gapext 1.0  
 arched: 2185239 seqs, 112599159 residues  
 tal number of hits satisfying chosen parameters: 4370478  
 nimum DB seq length: 0  
 ximum DB seq length: 2000000000  
 st-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : N Geneseq\_101002.\*  
 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
 24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39.4	96.1	143068	21	AAF21105 Human low adenosin
2	39.4	96.1	143068	21	AAF21272 Human low adenosin
3	39.4	96.1	143068	21	AAA34983 Human adenosine re
4	39.4	96.1	143068	21	AAA35150 Human adenosine re
5	39.4	96.1	143068	24	ABL68124 Ovary cancer relat
6	39.4	96.1	149412	21	AAA35151 Human adenosine re
7	39.4	96.1	152740	21	AAF21273 Human low adenosin
8	25.6	62.4	1624	22	ABAI5771 Human immune/haema
9	25.6	62.4	1624	22	AAK73388

c 10	24.2	59.0	633	22	AAJ14736 Human breast cance
c 11	24.2	59.0	908	23	AAJ84558 DNA encoding novel
c 12	24	58.5	360	22	AAI82165 Human polynucleoti
c 13	24	58.5	1001	22	AAI81867 Human polynucleoti
c 14	24	58.5	1311	22	AAH94457 Human foetal cDNA,
c 15	24	58.5	3565	16	AAQ84051 Sequence encoding
c 16	24	58.5	8728	22	AAK87277 Human immune/haema
c 17	24	58.5	10528	22	ABL32334 Human immune syste
c 18	24	58.5	17335	23	ABK42393 Genomic sequence #
c 19	24	58.5	19882	23	ABK42394 Genomic sequence #
c 20	24	58.5	21982	22	AAK67644 Human immune/haema
c 21	23.6	57.6	604	22	AAJ20925 Human breast cance
c 22	23.6	57.6	604	22	AAJ23598 Human breast cance
c 23	23.4	57.1	669	24	ABL89767 Human polynucleoti
c 24	23.4	57.1	38374	24	ABN96966 Gene #3464 used to
c 25	23.4	57.1	38374	24	ABL68363 Kidney cancer rela
c 26	23.4	57.1	38374	24	ABL68364 Kidney cancer rela
c 27	23.4	57.1	38374	24	ABL68824 Kidney cancer rela
c 28	23.4	57.1	1503900	22	AAK95240 Human neuregulin-1
c 29	23.4	57.1	1503900	22	AAK96733 Human polynucleoti
c 30	23.2	56.6	496	22	AAI86106 Human reproductive
c 31	23.2	56.6	8835	22	AAJ02806 Human cDNA clone (
c 32	23	56.1	605	22	AAH07386 Human immune/haema
c 33	22.6	55.1	47090	22	AAK68725 Human immune/haema
c 34	22.6	55.1	47090	22	AAK78219 Human immune/haema
c 35	22.6	55.1	75384	22	AAK85590 Human immune/haema
c 36	22.6	55.1	145831	24	ABL62309 Colon adenocarcino
c 37	22.6	55.1	145831	24	ABL66806 Lung cancer relate
c 38	22.6	55.1	145831	24	ABL68588 Kidney cancer rela
c 39	22.6	55.1	145831	24	ABL69213 Prostate cancer re
c 40	22.4	54.6	322	22	AAJ36435 Human musculoskele
c 41	22.4	54.6	329	22	AAJ31783 Human polynucleoti
c 42	22.4	54.6	349	22	ABAJ4274 Human nervous syst
c 43	22.4	54.6	371	22	AAJ09563 Human breast cance
c 44	22.4	54.6	379	22	AAJ17455 Human breast cance
c 45	22.4	54.6	387	22	AAK67876 Human immune/haema

## ALIGNMENTS

RESULT 1  
 AAF21105  
 ID AAF21105 standard; DNA; 143068 BP.  
 XX AAF21105;  
 AC AAF21105;  
 XX AAF21105;  
 DT 14-MAR-2001 (first entry)  
 XX Human low adenosine antisense oligonucleotide related sequence #2672.  
 XX Low adenosine antisense oligonucleotide; phosphorothioate; allergy;  
 KW human; airway disorder; bronchoconstriction; lung inflammation;  
 KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;  
 KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;  
 KW respiratory obstruction; pulmonary obstruction; impeded respiration;  
 KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;  
 KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;  
 KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;  
 KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;  
 KW cancer; ss.  
 XX Homo sapiens.  
 OS WO2000062736-A2.  
 XX 26-OCT-2000.  
 PD 24-MAR-2000; 2000WO-US08020.  
 PF 06-APR-1999; 99US-0127958.  
 XX (UYEC-) UNIV EAST CAROLINA.  
 PA



(NYCE/) NYCE J W.  
 Nyce JW;  
 WPI; 2000-679539/66.  
 Low adenosine (A) content antisense oligonucleotides which do not trigger adenosine receptors during metabolism, useful e.g. for treating cancers and respiratory obstructions -  
 Disclosure; Page 924-957; 1592pp; English.  
 The present invention describes low adenosine (A) content antisense oligonucleotides and compositions (I) comprising them. In the antisense oligonucleotides the A is replaced by a 'Universal' or alternative base. (I) can have respiratory, bronchodilator, antiinflammatory, analgesic, immunosuppressive, antiasthmatic, hypotensive and cytostatic activities. The antisense oligonucleotides and (I) can be used to down-regulate the expression and/or activity of target polypeptides associated with lung/respiratory disorders and malignancies, such as stimulating and activating peptide factors and transmitters, transcription factors, immunoglobulins and antibodies, antibody receptors, cytokines and chemokines, endogenously produced specific and non-specific enzymes, binding proteins, adhesion molecules and their receptors, cytokine and chemokine receptors, adenosine receptors, bradykinin receptors, central nervous system (CNS) and peripheral nervous and non-nervous system receptors, CNS and peripheral nervous and non-nervous system transmitters, defensins, growth factors, vasoactive peptides and receptors, binding proteins and malignancy associated proteins. The antisense oligonucleotides may be used in this way to treat disorders including respiratory obstruction (especially pulmonary obstruction and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or surfactant hypoproduction which are associated with a disease or condition selected from pulmonary vasoconstriction, inflammation, allergies, asthma, impeded respiration, respiratory distress syndrome (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), pulmonary transplantation rejection, pulmonary infections, bronchitis, and/or cancer. AAF18434 to AAF21543 represent human polynucleotide fragments and antisense oligonucleotides used in the exemplification of the present invention.  
 Sequence 143068 BP; 41194 A; 30122 C; 32403 G; 39349 T; 0 other;  
 Query Match 96.1%; Score 39.4; DB 21; Length 143068;  
 Best Local Similarity 97.6%; Pred. No. 1.1e-05;  
 Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 1 TCAGTTGCTGAGAGCTGGAATACCGAGCTGCTGAGAC 41  
 42703 TCAGTTGCTGAGAGCTGGAATACCGAGCTGCTGAGAC 42743  
 RESULT 2  
 AAF21272  
 AAF21272 standard; DNA; 143068 BP.  
 AAF21272;  
 14-MAR-2001 (first entry)  
 Human low adenosine antisense oligonucleotide related sequence #2839.  
 Low adenosine antisense oligonucleotide; phosphorothioate; allergy; human; airway disorder; bronchoconstriction; lung inflammation; surfactant depletion; respiratory; bronchodilator; antiinflammatory; immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic; respiratory obstruction; pulmonary obstruction; impeded respiration; surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS; respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis; pulmonary hypertension; emphysema; pulmonary transplantation rejection; chronic obstructive pulmonary disease; pulmonary infection; bronchitis; cancer; SB.

XX OS Homo sapiens.  
 XX PN WO200062736-A2.  
 XX PD 26-OCT-2000.  
 XX PF 24-MAR-2000; 2000WO-US08020.  
 XX PR 06-APR-1999; 99US-0127958.  
 XX PA (UYEC-) UNIV EAST CAROLINA.  
 XX PY (NYCE/) NYCE J W.  
 XX PI Nyce JW;  
 XX DR WPI; 2000-679539/66.  
 XX PT Low adenosine (A) content antisense oligonucleotides which do not trigger adenosine receptors during metabolism, useful e.g. for treating cancers and respiratory obstructions -  
 XX PT Disclosure; Page 1186-1219; 1592pp; English.  
 XX PS The present invention describes low adenosine (A) content antisense oligonucleotides and compositions (I) comprising them. In the antisense oligonucleotides the A is replaced by a 'Universal' or alternative base. (I) can have respiratory, bronchodilator, antiinflammatory, analgesic, immunosuppressive, antiasthmatic, hypotensive and cytostatic activities. The antisense oligonucleotides and (I) can be used to down-regulate the expression and/or activity of target polypeptides associated with lung/respiratory disorders and malignancies, such as stimulating and activating peptide factors and transmitters, transcription factors, immunoglobulins and antibodies, antibody receptors, cytokines and chemokines, endogenously produced specific and non-specific enzymes, binding proteins, adhesion molecules and their receptors, cytokine and chemokine receptors, adenosine receptors, bradykinin receptors, central nervous system (CNS) and peripheral nervous and non-nervous system receptors, CNS and peripheral nervous and non-nervous system transmitters, defensins, growth factors, vasoactive peptides and receptors, binding proteins and malignancy associated proteins. The antisense oligonucleotides may be used in this way to treat disorders including respiratory obstruction (especially pulmonary obstruction and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or surfactant hypoproduction which are associated with a disease or condition selected from pulmonary vasoconstriction, inflammation, allergies, asthma, impeded respiration, respiratory distress syndrome (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), pulmonary transplantation rejection, pulmonary infections, bronchitis, and/or cancer. AAF18434 to AAF21543 represent human polynucleotide fragments and antisense oligonucleotides used in the exemplification of the present invention.  
 XX SQ Sequence 143068 BP; 41194 A; 30122 C; 32403 G; 39349 T; 0 other;  
 Query Match 96.1%; Score 39.4; DB 21; Length 143068;  
 Best Local Similarity 97.6%; Pred. No. 1.1e-05;  
 Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 TCAGTTGCTGAGAGCTGGAATACCGAGCTGCTGAGAC 41  
 DB 42703 TCAGTTGCTGAGAGCTGGAATACCGAGCTGCTGAGAC 42743  
 RESULT 3  
 AAA34983  
 ID AAA34983 standard; DNA; 143068 BP.  
 XX AC AAA34983;  
 XX DT 28-JUL-2000 (first entry)  
 XX

Human adenosine receptor related polynucleotide SEQ ID NO:2672.

Human; adenosine receptor; low adenosine antisense oligonucleotide; phosphorothioate; impaired respiration; inflammation; allergy; allergic disease; bronchoconstriction; inhibitor; antiinflammatory; antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway; lung disease; ischaemic condition; pulmonary vasoconstriction; asthma; respiratory distress syndrome; pain; cystic fibrosis; emphysema; pulmonary hypertension; chronic obstructive pulmonary disease; COPD; cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.

Homo sapiens.

WO200009525-A2.

24-FEB-2000.

03-AUG-1999; 99WO-US17712.

03-AUG-1998; 98US-0095212.

(UYEC-) UNIV EAST CAROLINA.

Nyce JW;

WPI; 2000-205971/18.

New antisense oligonucleotides useful for treating e.g. pulmonary vasoconstriction, inflammation, allergies, asthma, hypertension, bronchitis, emphysema, respiratory distress syndrome, ischemia or cancers -

Disclosure; Page 851-882; 1343pp; English.

The present invention describes a new composition comprising an antisense oligonucleotide (ON) with low adenosine (up to 15%), which targets nucleic acids involved in bronchoconstriction, allergies, and/or inflammation. The ON can have antiinflammatory, antiallergic, antiasthmatic, cytostatic and analgesic activities. The composition is useful for the treatment of diseases associated with inflammation, impaired airways, including lung disease and diseases whose secondary effects afflict the lungs of a subject. They can be used for treating e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma, impaired respiration, respiratory distress syndrome, pain, cystic fibrosis, pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), and cancers such as leukaemias, lymphomas, carcinomas, and cancers which may metastasize to the lungs, including breast and prostate cancer. The reduction of the adenosine content of the ONs reduces side effects. The A-containing ONs break down with the release of deoxyadenosine which activates adenosine receptors causing bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the nucleotide sequences given in the sequence listing from the present invention, which correspond to SEQ ID NO:1 to 2815, and then the last 185 sequences are also called SEQ ID NO:1 to 185, but the sequences differ from the previously named sequences. SEQ ID NO:11 to 1680 (AAA32323 to AAA33992) are specifically claimed ONs from the present invention. N.B. Sequences given in the disclosure of the present invention do not match up with their corresponding SEQ ID NO: sequences given in the sequence listing.

Sequence 143068 BP; 41194 A; 30122 C; 32402 G; 39350 T; 0 other;

Query Match 96.1%; Score 39.4; DB 21; Length 143068;  
Best Local Similarity 97.6%; Pred. No. 1.1e-05;  
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 TCAGTGTCTGAGAGCCTGAATACAGGACTGCTGAGAC 41

42703 TCAGTGTCTGAGAGCCTGACATACAGGACTGCTGAGAC 42743

SULT 4  
A35150

ID AAA35150 standard; DNA; 143068 BP.  
XX AAA35150;  
AC AAA35150;  
XX  
DT 28-JUL-2000 (first entry)  
XX  
DE Human adenosine receptor related polynucleotide 2nd SEQ ID NO:24.  
XX  
XX Human; adenosine receptor; low adenosine antisense oligonucleotide; phosphorothioate; impaired respiration; inflammation; allergy; allergic disease; bronchoconstriction; inhibitor; antiinflammatory; antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway; lung disease; ischaemic condition; pulmonary vasoconstriction; asthma; respiratory distress syndrome; pain; cystic fibrosis; emphysema; pulmonary hypertension; chronic obstructive pulmonary disease; COPD; cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200009525-A2.  
XX  
XX 24-FEB-2000.  
XX  
XX 03-AUG-1999; 99WO-US17712.  
XX  
XX 03-AUG-1998; 98US-0095212.  
XX  
XX (UYEC-) UNIV EAST CAROLINA.  
XX  
XX Nyce JW;  
XX  
XX WPI; 2000-205971/18.  
XX  
XX New antisense oligonucleotides useful for treating e.g. pulmonary vasoconstriction, inflammation, allergies, asthma, hypertension, bronchitis, emphysema, respiratory distress syndrome, ischemia or cancers -  
XX  
XX Disclosure; Page 1106-1138; 1343pp; English.  
XX  
XX The present invention describes a new composition comprising an antisense oligonucleotide (ON) with low adenosine (up to 15%), which targets nucleic acids involved in bronchoconstriction, allergies, and/or inflammation. The ON can have antiinflammatory, antiallergic, antiasthmatic, cytostatic and analgesic activities. The compositions are useful for the treatment of diseases associated with inflammation, impaired airways, including lung disease and diseases whose secondary effects afflict the lungs of a subject. They can be used for treating e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma, impaired respiration, respiratory distress syndrome, pain, cystic fibrosis, pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), and cancers such as leukaemias, lymphomas, carcinomas, and cancers which may metastasize to the lungs, including breast and prostate cancer. The reduction of the adenosine content of the ONs reduces side effects. The A-containing ONs break down with the release of deoxyadenosine which activates adenosine receptors causing bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the nucleotide sequences given in the sequence listing from the present invention, which correspond to SEQ ID NO:1 to 2815, and then the last 185 sequences are also called SEQ ID NO:1 to 185, but the sequences differ from the previously named sequences. SEQ ID NO:11 to 1680 (AAA32323 to AAA33992) are specifically claimed ONs from the present invention. N.B. Sequences given in the disclosure of the present invention do not match up with their corresponding SEQ ID NO: sequences given in the sequence listing.  
XX  
SQ Sequence 143068 BP; 41194 A; 30126 C; 32402 G; 39346 T; 0 other;  
Query Match 96.1%; Score 39.4; DB 21; Length 143068;  
Best Local Similarity 97.6%; Pred. No. 1.1e-05;  
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TCAGTGTCTGAGAGCCTGAATACAGGACTGCTGAGAC 41

42703 TCAGTGTCTGAGAGCTGCATACACAGGACTGCCTGAGAC 42743

RESULT 5  
168124  
ABL68124 standard; DNA; 143068 BP.

ABL68124;

15-MAY-2002 (first entry)

Ovary cancer related gene sequence SEQ ID NO:6461.

Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
gene; ds.

Homo sapiens.

WO200194629-A2.

13-DEC-2001.

30-MAY-2001; 2001WO-US10638.

05-JUN-2000; 2000US-209473P.

05-JUN-2000; 2000US-209531P.

18-SEP-2000; 2000US-231133P.

18-SEP-2000; 2000US-231617P.

20-SEP-2000; 2000US-234009P.

20-SEP-2000; 2000US-234034P.

20-SEP-2000; 2000US-234052P.

22-SEP-2000; 2000US-234509P.

22-SEP-2000; 2000US-234567P.

25-SEP-2000; 2000US-234923P.

25-SEP-2000; 2000US-234924P.

25-SEP-2000; 2000US-235077P.

25-SEP-2000; 2000US-235082P.

25-SEP-2000; 2000US-235134P.

25-SEP-2000; 2000US-235280P.

26-SEP-2000; 2000US-235637P.

26-SEP-2000; 2000US-235638P.

27-SEP-2000; 2000US-235711P.

27-SEP-2000; 2000US-235720P.

27-SEP-2000; 2000US-235840P.

27-SEP-2000; 2000US-235863P.

28-SEP-2000; 2000US-236028P.

28-SEP-2000; 2000US-236032P.

28-SEP-2000; 2000US-236033P.

28-SEP-2000; 2000US-236034P.

28-SEP-2000; 2000US-236109P.

28-SEP-2000; 2000US-236111P.

29-SEP-2000; 2000US-236842P.

29-SEP-2000; 2000US-236891P.

02-OCT-2000; 2000US-237172P.

02-OCT-2000; 2000US-237173P.

02-OCT-2000; 2000US-237278P.

02-OCT-2000; 2000US-237294P.

02-OCT-2000; 2000US-237295P.

03-OCT-2000; 2000US-237316P.

03-OCT-2000; 2000US-237425P.

03-OCT-2000; 2000US-237598P.

03-OCT-2000; 2000US-237604P.

03-OCT-2000; 2000US-237606P.

03-OCT-2000; 2000US-237608P.

01-NOV-2000; 2000US-244867P.

01-NOV-2000; 2000US-245084P.

(AVAL-) AVALON PHARM.

Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;

PI Soppet DR, Weaver Z;  
XX WPI; 2002-188264/24.  
XX  
XX  
XX Screening for anti-neoplastic agent involves exposing cells to a  
PT chemical agent to be tested for anti-neoplastic activity, and  
PT determining a change in expression of a gene of a signature gene set -  
XX  
XX Claim 1; SEQ ID 6461; 44pp; English.  
XX  
XX The present invention describes a method (M1) for screening for an  
CC anti-neoplastic agent. The method involves exposing cells to a chemical  
CC agent to be tested for anti-neoplastic activity, determining a change in  
CC expression of at least one gene (I) of a signature gene set, where (I)  
CC comprises a sequence (S) selected from 8447 sequences (given in ABU61664  
CC to ABU70110), or is at least 95% identical to (S), where a change in  
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
CC activity and can be used in gene therapy. M1 can be used for screening  
CC an anti-neoplastic agent, and can be used for producing a product which  
CC is the data collected with respect to the anti-neoplastic agent as a  
CC result of M1, and the data is sufficient to convey the chemical  
CC structure and/or properties of the agent. M1 can be used in the  
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,  
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
CC carcinoma, papillary carcinoma and Wilm's tumour.  
XX  
XX Sequence 143068 BP; 41194 A; 30122 C; 32403 G; 39349 T; 0 other;  
SQ  
Query Match 96.1%; Score 39.4; DB 24; Length 143068;  
Best Local Similarity 97.6%; Pred. No. 1.1e-05;  
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TCAGTGTCTGAGAGCTGCATACACAGGACTGCCTGAGAC 41  
DB 42703 TCAGTGTCTGAGAGCTGCATACACAGGACTGCCTGAGAC 42743  
RESULT 6  
AA35151  
ID AAA35151 standard; DNA; 149412 BP.  
XX  
XX AAA35151;  
XX  
XX 28-JUL-2000 (first entry)  
XX  
XX Human adenosine receptor related polynucleotide 2nd SEQ ID NO:25.  
XX  
XX Human; adenosine receptor; low adenosine antisense oligonucleotide;  
XX phosphorothioate; impaired respiration; inflammation; allergy;  
XX allergic disease; bronchoconstriction; inhibitor; antiinflammatory;  
XX antiallergic; antilasthmatic; cytostatic; analgesic; impaired airway;  
XX lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;  
XX respiratory distress syndrome; pain; cystic fibrosis; emphysema;  
XX pulmonary hypertension; chronic obstructive pulmonary disease; COPD;  
XX cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200009525-A2.  
XX  
XX 24-FEB-2000.  
XX  
XX 03-AUG-1999; 99WO-US17712.  
XX  
XX 03-AUG-1998; 98US-0095212.  
XX  
XX (UYEC-) UNIV EAST CAROLINA.  
XX  
XX Nyce JW;  
XX  
XX WPI; 2000-205971/18.  
XX  
XX

New antisense oligonucleotides useful for treating e.g. pulmonary vasoconstriction, inflammation, allergies, asthma, hypertension, bronchitis, emphysema, respiratory distress syndrome, ischemia or cancers -

Disclosure; Page 1138-1171; 1343pp; English.

The present invention describes a new composition comprising an antisense oligonucleotide (ON) with low adenine (up to 15%), which targets nucleic acids involved in bronchoconstriction, allergies, and/or inflammation. The ON can have antiinflammatory, antiallergic, antiasthmatic, cytostatic and analgesic activities. The compositions are useful for the treatment of diseases associated with inflammation, impaired airways, including lung disease and diseases whose secondary effects affect the lungs of a subject. They can be used for treating e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma, impaired respiration, respiratory distress syndrome, pain, cystic fibrosis, pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), and cancers such as leukaemias, lymphomas, carcinomas, and cancers which may metastasize to the lungs, including breast and prostate cancer. The reduction of the adenine content of the ONs reduces side effects. The A-containing ONs break down with the release of deoxyadenosine which activates adenosine receptors causing bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the nucleotide sequences given in the sequence listing from the present invention, which correspond to SEQ ID NO:1 to 2815, and then the last 185 sequences are also called SEQ ID NO:1 to 185, but the sequences differ from the previously named sequences. SEQ ID NO:11 to 1680 (AAA32323 to AAA33992) are specifically claimed ONs from the present invention. N.B. Sequences given in the disclosure of the present invention do not match up with their corresponding SEQ ID NO: sequences given in the sequence listing.

Sequence 149412 BP; 43049 A; 31388 C; 33852 G; 41123 T; 0 other;

Query Match 96.1%; Score 39.4; DB 21; Length 149412;  
Best Local Similarity 97.6%; Pred. No. 1.1e-05;  
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 TCAGTTCGTGAGAGCTGAAATACACAGACTGCTGAGAC 41

49047 TCAGTTCGTGAGAGCTGACATACACAGACTGCTGAGAC 49087

RESULT 7

AAF21273 standard; DNA; 152740 BP.

AAF21273;

14-MAR-2001 (first entry)

Human low adenine antisense oligonucleotide related sequence #2840.

Low adenine antisense oligonucleotide; phosphorothioate; allergy; human; airway disorder; bronchoconstriction; lung inflammation; surfactant depletion; respiratory; bronchodilator; antiinflammatory; immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic; respiratory obstruction; pulmonary obstruction; impeded respiration; surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS; respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis; pulmonary hypertension; emphysema; pulmonary transplantation rejection; chronic obstructive pulmonary disease; pulmonary infection; bronchitis; cancer; ss.

Homo sapiens.

WO200062736-A2.

26-OCT-2000.

24-MAR-2000; 2000WO-US08020.

XX

06-APR-1999; 99US-0127958.

XX

(UYEC-) UNIV EAST CAROLINA.

PA

(NYCE/) NYCE J W.

XX

NYce JW;

PI

WPI; 2000-679539/66.

XX

Low adenine (A) content antisense oligonucleotides which do not trigger adenosine receptors during metabolism, useful e.g. for treating cancers and respiratory obstructions -

Disclosure; Page 1219-1254; 1592pp; English.

The present invention describes low adenine (A) content antisense oligonucleotides and compositions (I) comprising them. In the antisense oligonucleotides the A is replaced by a 'Universal' or alternative base. (I) can have respiratory, bronchodilator, antiinflammatory, analgesic, immunosuppressive, antiasthmatic, hypotensive and cytostatic activities. The antisense oligonucleotides and (I) can be used to down-regulate the expression and/or activity of target polypeptides associated with lung/respiratory disorders and malignancies, such as stimulating and activating peptide factors and transmitters, transcription factors, immunoglobulins and antibodies, antibody receptors, cytokines and chemokines, endogenously produced specific and non-specific enzymes, binding proteins, adhesion molecules and their receptors, cytokine and chemokine receptors, adenosine receptors, bradykinin receptors, central nervous system (CNS) and peripheral nervous and non-nervous system receptors, CNS and peripheral nervous and non-nervous system peptide transmitters, defensins, growth factors, vasoactive peptides and receptors, binding proteins and malignancy associated proteins. The antisense oligonucleotides may be used in this way to treat disorders including respiratory obstruction (especially pulmonary obstruction and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or surfactant hypoproduction which are associated with a disease or condition selected from pulmonary vasoconstriction, inflammation, allergies, asthma, impeded respiration, respiratory distress syndrome (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), pulmonary transplantation rejection, pulmonary infections, bronchitis, and/or cancer. AAF18434 to AAF21543 represent human polynucleotide fragments and antisense oligonucleotides used in the exemplification of the present invention.

Sequence 152740 BP; 44169 A; 32023 C; 34549 G; 41939 T; 0 other;

Query Match 96.1%; Score 39.4; DB 21; Length 152740;  
Best Local Similarity 97.6%; Pred. No. 1.1e-05;  
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 TCAGTTCGTGAGAGCTGAAATACACAGACTGCTGAGAC 41

49047 TCAGTTCGTGAGAGCTGACATACACAGACTGCTGAGAC 49087

RESULT 8

ABAA15771/c

ID ABAA15771 standard; DNA; 1624 BP.

XX

ABAA15771;

XX

23-JAN-2002 (first entry)

XX

Human nervous system related polynucleotide SEQ ID NO 8102.

XX

Human; nootropic; neuroprotective; cytostatic; dermatological; virucide; immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary; antiparkinsonian; antiskinning; antianemic; antiarthritic; cancer; antirheumatic; hepatotropic; cerebroprotective; antiinflammatory; antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal; antiparasitic; cardiant; immune disorder; cardiovascular disorder;

neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.

Homo sapiens.

WO200159063-A2.

16-AUG-2001.

17-JAN-2001; 2001WO-US01334.

31-JAN-2000; 2000US-0179065.

04-FEB-2000; 2000US-0180628.

24-FEB-2000; 2000US-0184664.

02-MAR-2000; 2000US-0186350.

16-MAR-2000; 2000US-0189874.

17-MAR-2000; 2000US-0190076.

18-APR-2000; 2000US-0198123.

19-MAY-2000; 2000US-0205515.

07-JUN-2000; 2000US-0209467.

28-JUN-2000; 2000US-0214886.

30-JUN-2000; 2000US-0215135.

07-JUL-2000; 2000US-0216647.

07-JUL-2000; 2000US-0216980.

11-JUL-2000; 2000US-0217487.

11-JUL-2000; 2000US-0217496.

14-JUL-2000; 2000US-0218290.

26-JUL-2000; 2000US-0220963.

26-JUL-2000; 2000US-0220964.

14-AUG-2000; 2000US-0224518.

14-AUG-2000; 2000US-0224519.

14-AUG-2000; 2000US-0225213.

14-AUG-2000; 2000US-0225214.

14-AUG-2000; 2000US-0225266.

14-AUG-2000; 2000US-0225267.

14-AUG-2000; 2000US-0225268.

14-AUG-2000; 2000US-0225270.

14-AUG-2000; 2000US-0225447.

14-AUG-2000; 2000US-0225757.

14-AUG-2000; 2000US-0225758.

14-AUG-2000; 2000US-0225759.

18-AUG-2000; 2000US-0226279.

22-AUG-2000; 2000US-0226681.

22-AUG-2000; 2000US-0226868.

22-AUG-2000; 2000US-0227182.

23-AUG-2000; 2000US-0227009.

PR	25-SEP-2000;	2000US-0234998.
PR	26-SEP-2000;	2000US-0235484.
PR	27-SEP-2000;	2000US-0235834.
PR	27-SEP-2000;	2000US-0235836.
PR	29-SEP-2000;	2000US-0236327.
PR	29-SEP-2000;	2000US-0236367.
PR	29-SEP-2000;	2000US-0236368.
PR	29-SEP-2000;	2000US-0236369.
PR	29-SEP-2000;	2000US-0236370.
PR	02-OCT-2000;	2000US-0236802.
PR	02-OCT-2000;	2000US-0237037.
PR	02-OCT-2000;	2000US-0237038.
PR	02-OCT-2000;	2000US-0237039.
PR	02-OCT-2000;	2000US-0237040.
PR	13-OCT-2000;	2000US-0239335.
PR	13-OCT-2000;	2000US-0239337.
PR	20-OCT-2000;	2000US-0240960.
PR	20-OCT-2000;	2000US-0241785.
PR	20-OCT-2000;	2000US-0241786.
PR	20-OCT-2000;	2000US-0241787.
PR	20-OCT-2000;	2000US-0241808.
PR	20-OCT-2000;	2000US-0241809.
PR	20-OCT-2000;	2000US-0241826.
PR	20-OCT-2000;	2000US-0242221.
PR	01-NOV-2000;	2000US-0244617.
PR	08-NOV-2000;	2000US-0246474.
PR	08-NOV-2000;	2000US-0246475.
PR	08-NOV-2000;	2000US-0246476.
PR	08-NOV-2000;	2000US-0246477.
PR	08-NOV-2000;	2000US-0246478.
PR	08-NOV-2000;	2000US-0246523.
PR	08-NOV-2000;	2000US-0246524.
PR	08-NOV-2000;	2000US-0246525.
PR	08-NOV-2000;	2000US-0246526.
PR	08-NOV-2000;	2000US-0246527.
PR	08-NOV-2000;	2000US-0246528.
PR	08-NOV-2000;	2000US-0246532.
PR	08-NOV-2000;	2000US-0246609.
PR	08-NOV-2000;	2000US-0246610.
PR	08-NOV-2000;	2000US-0246611.
PR	08-NOV-2000;	2000US-0246613.
PR	17-NOV-2000;	2000US-0249207.
PR	17-NOV-2000;	2000US-0249208.
PR	17-NOV-2000;	2000US-0249209.
PR	17-NOV-2000;	2000US-0249210.
PR	17-NOV-2000;	2000US-0249211.
PR	17-NOV-2000;	2000US-0249212.
PR	17-NOV-2000;	2000US-0249213.
PR	17-NOV-2000;	2000US-0249214.
PR	17-NOV-2000;	2000US-0249215.
PR	17-NOV-2000;	2000US-0249216.
PR	17-NOV-2000;	2000US-0249217.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249244.
PR	17-NOV-2000;	2000US-0249245.
PR	17-NOV-2000;	2000US-0249264.
PR	17-NOV-2000;	2000US-0249265.
PR	17-NOV-2000;	2000US-0249297.
PR	17-NOV-2000;	2000US-0249299.
PR	17-NOV-2000;	2000US-0249300.
PR	01-DEC-2000;	2000US-0250391.
PR	01-DEC-2000;	2000US-0251160.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251989.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259678.



20-OCT-2000; 2000US-0241786.  
 20-OCT-2000; 2000US-0241787.  
 20-OCT-2000; 2000US-0241808.  
 20-OCT-2000; 2000US-0241809.  
 20-OCT-2000; 2000US-0241826.  
 01-NOV-2000; 2000US-0244617.  
 08-NOV-2000; 2000US-0246474.  
 08-NOV-2000; 2000US-0246475.  
 08-NOV-2000; 2000US-0246476.  
 08-NOV-2000; 2000US-0246477.  
 08-NOV-2000; 2000US-0246478.  
 08-NOV-2000; 2000US-0246523.  
 08-NOV-2000; 2000US-0246524.  
 08-NOV-2000; 2000US-0246525.  
 08-NOV-2000; 2000US-0246526.  
 08-NOV-2000; 2000US-0246527.  
 08-NOV-2000; 2000US-0246528.  
 08-NOV-2000; 2000US-0246532.  
 08-NOV-2000; 2000US-0246509.  
 08-NOV-2000; 2000US-0246610.  
 08-NOV-2000; 2000US-0246611.  
 08-NOV-2000; 2000US-0246613.  
 17-NOV-2000; 2000US-0249207.  
 17-NOV-2000; 2000US-0249208.  
 17-NOV-2000; 2000US-0249209.  
 17-NOV-2000; 2000US-0249210.  
 17-NOV-2000; 2000US-0249211.  
 17-NOV-2000; 2000US-0249212.  
 17-NOV-2000; 2000US-0249213.  
 17-NOV-2000; 2000US-0249214.  
 17-NOV-2000; 2000US-0249215.  
 17-NOV-2000; 2000US-0249216.  
 17-NOV-2000; 2000US-0249217.  
 17-NOV-2000; 2000US-0249218.  
 17-NOV-2000; 2000US-0249244.  
 17-NOV-2000; 2000US-0249245.  
 17-NOV-2000; 2000US-0249264.  
 17-NOV-2000; 2000US-0249265.  
 17-NOV-2000; 2000US-0249297.  
 17-NOV-2000; 2000US-0249299.  
 17-NOV-2000; 2000US-0249300.  
 01-DEC-2000; 2000US-0250160.  
 01-DEC-2000; 2000US-0250391.  
 05-DEC-2000; 2000US-0251030.  
 05-DEC-2000; 2000US-0251988.  
 05-DEC-2000; 2000US-0256719.  
 06-DEC-2000; 2000US-0251479.  
 08-DEC-2000; 2000US-0251856.  
 08-DEC-2000; 2000US-0251868.  
 08-DEC-2000; 2000US-0251869.  
 08-DEC-2000; 2000US-0251989.  
 08-DEC-2000; 2000US-0251990.  
 11-DEC-2000; 2000US-0254097.  
 05-JAN-2001; 2001US-0259678.  
 (HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483426/52.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis -

Disclosure; SEQ ID NO 28200; 3071pp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I) amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased

CC expression by rectifying mutations or deletions in a patient's genome  
 CC that affect the activity of (I) by expressing inactive proteins or to  
 CC supplement the patients own production of (I). Additionally, (I)  
 CC polynucleotides may be used to produce the secreted (I), by inserting the  
 CC the nucleic acids into a host cell and culturing the cell to express the  
 CC protein. (I) proteins and polynucleotides may be used to prevent,  
 CC diagnose and treat immune/haematopoietic-related diseases, especially  
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
 CC to AAK87694 represent human immune/haematopoietic antigen genomic  
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
 CC represent sequences used in the exemplification of the present invention.  
 XX  
 SQ Sequence 1624 BP; 438 A; 393 C; 350 G; 443 T; 0 other;

Query Match 62.4%; Score 25.6; DB 22; Length 1624;  
 Best Local Similarity 77.5%; Pred. No. 1.5;  
 Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 CAGTTGCTGAGAAGCCTGAATACACGAGCTGCTGAGAC 41  
 |||||  
 Db 706 CAGTACTCAGAGCCTGAGTGGAGAGCTGCTTGAGCC 745

RESULT 10

AAI14736  
 ID AAI14736 standard; cDNA; 633 BP.

XX AAI14736;

XX 07-DEC-2001 (first entry)

XX Human breast cancer expressed polynucleotide 7193.

XX Human; breast cancer; cell marker; cytostatic; ss.

XX Homo sapiens.

XX WO200151628-A2.

XX 19-JUL-2001.

XX 10-JAN-2001; 2001WO-US00798.

XX 14-JAN-2000; 2000US-0176077.

XX 14-MAR-2000; 2000US-0189167.

XX 24-MAR-2000; 2000US-0192099.

XX 29-MAR-2000; 2000US-0193480.

XX 15-MAY-2000; 2000US-0205230.

XX 09-JUN-2000; 2000US-0211315.

XX 25-JUL-2000; 2000US-0220534.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Lillie J, Xu Y, Wang Y, Steinmann K;

XX WPI; 2001-451856/48.

XX New peptide useful as a marker for the diagnosis of breast cancer -

XX Claim 1; Page 1300; 3695pp; English.

XX The invention relates to human breast cancer expressed polynucleotides  
 CC (AAI07544-AAI26789) and methods of assessing whether a patient is  
 CC afflicted with breast cancer by examining the correlation between the  
 CC expression of certain markers and the cancerous state of breast cells.  
 CC The polynucleotides and encoded polypeptides are potential markers for  
 CC detecting, diagnosing, monitoring, characterising treating and  
 CC potentially preventing breast cancer. The polynucleotides and encoded  
 CC polypeptides are also useful for isolating compounds with cytostatic  
 CC activity.

XX Sequence 633 BP; 223 A; 92 C; 178 G; 136 T; 4 other;

Query Match 59.0%; Score 24.2; DB 22; Length 633;  
Best Local Similarity 76.3%; Pred. No. 4.4;  
Matches 29; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

1 TCAGTCTCAGAGCCTGAATACAGGACTGCTGAG 38  
|||||  
522 TCAGTCTCAGAGCCTGACGTANGAGAAATGCTTGA 559

ULT 11  
384558/c  
AAS84558 standard; cDNA; 908 BP.

AAS84558;

13-FEB-2002 (first entry)

DNA encoding novel human diagnostic protein #20362.

Human; chromosome mapping; gene mapping; gene therapy; forensic;  
food supplement; medical imaging; diagnostic; genetic disorder; ss.

Homo sapiens.

WO200175067-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-US08631.

31-MAR-2000; 2000US-0540217.

23-AUG-2000; 2000US-0649167.

(HYSE-) HYSEQ INC.

Drmanac RT, Liu C, Tang YT;

WPI; 2001-639362/73.

P-PSDB; ABG20371.

New isolated polynucleotide and encoded polypeptides, useful in  
diagnostics, forensics, gene mapping, identification of mutations  
responsible for genetic disorders or other traits and to assess  
biodiversity -

Claim 1; SEQ ID No 20362; 1032p; English.

The invention relates to isolated polynucleotide (I) and  
polypeptide (II) sequences. (I) is useful as hybridisation probes,  
polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
and gene mapping, and in recombinant production of (II). The  
polynucleotides are also used in diagnostics as expressed sequence tags  
for identifying expressed genes. (I) is useful in gene therapy techniques  
to restore normal activity of (II) or to treat disease states involving  
(II). (II) is useful for generating antibodies against it, detecting or  
quantitating a polypeptide in tissue, as molecular weight markers and as  
a food supplement. (II) and its binding partners are useful for treating  
imaging of sites expressing (II). (I) and (II) are useful for treating  
disorders involving aberrant protein expression or biological activity.  
The polypeptide and polynucleotide sequences have applications in  
diagnostics, forensics, gene mapping, identification of mutations  
responsible for genetic disorders or other traits to assess biodiversity  
and to produce other types of data and products dependent on DNA and  
amino acid sequences. AAS64197-AAS94564 represent novel human  
diagnostic coding sequences of the invention.

Note: The sequence data for this patent did not appear in the printed  
specification, but was obtained in electronic format directly from WIPO  
at ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 908 BP; 265 A; 216 C; 210 G; 217 T; 0 other;

Query Match 59.0%; Score 24.2; DB 23; Length 908;  
Best Local Similarity 78.4%; Pred. No. 4.7;

Matches 29; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 3 AGTTCGTGAGAGCCTGAATACAGGACTGCTGAG 39  
|||||  
DB 820 AGTAGCGCAGCAGCCTGAATACAGGCTTCTCTAG 784

RESULT 12

AAI82165  
ID AAI82165 standard; cDNA; 360 BP.

XX AC AAI82165;

XX DT 06-NOV-2001 (first entry)

XX DE Human polynucleotide SEQ ID NO 2225.

XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorders; arthritis; inflammation; ss.

XX OS Homo sapiens.

XX PN WO200164835-A2.

XX PD 07-SEP-2001.

XX PF 26-FEB-2001; 2001WO-US04927.

XX PR 28-FEB-2000; 2000US-0515126.

XX PR 18-MAY-2000; 2000US-0577409.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-514838/56.

XX DR P-PSDB; AAO02234.

XX PT Isolated nucleic acids and polypeptides, useful for preventing  
PT diagnosing and treating e.g. leukaemia, inflammation and immune  
PT disorders -

XX PS Claim 1; SEQ ID NO 2225; 1399pp + Sequence Listing; English.

XX CC The invention relates to human polynucleotides (AAI79941-AAI93841) and  
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.

XX CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 360 BP; 110 A; 73 C; 93 G; 84 T; 0 other;

Query Match 58.5%; Score 24; DB 22; Length 360;

Best Local Similarity 75.0%; Pred. No. 4.7;

Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2 CAGTTCGTGAGAGCCTGAATACAGGACTGCTGAGAC 41

|||||  
DB 49 CAGCTACTCAGGAGGCTGAGTAGGAGAGCTGCTTGACC 88

RESULT 13



'I81867/C  
AAI81867 standard; cDNA; 1001 BP.  
AAI81867;  
06-NOV-2001 (first entry)  
Human polynucleotide SEQ ID NO 1927.  
Human; cytokine; cell proliferation; cell differentiation; gene therapy; vaccine; peptide therapy; stem cell growth factor; haematopoiesis; tissue growth factor; immunomodulatory; cancer; leukaemia; nervous system disorders; arthritis; inflammation; ss.  
Homo sapiens.  
WO200164835-A2.  
07-SEP-2001.  
26-FEB-2001; 2001WO-US04927.  
28-FEB-2000; 2000US-0515126.  
18-MAY-2000; 2000US-0577409.  
(HYSE-) HYSEQ INC.  
Tang YT, Liu C, Drmanac RT;  
WPI; 2001-514838/56.  
P-PSDB; AAO01936.  
Isolated nucleic acids and polypeptides, useful for preventing diagnosing and treating e.g. leukaemia, inflammation and immune disorders -  
Claim 1; SEQ ID NO 1927; 1399pp + Sequence Listing; English.  
The invention relates to human polynucleotides (AAI79941-AAI93841) and the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation.  
Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
Sequence 1001 BP; 258 A; 272 C; 239 G; 229 T; 3 other;  
Query Match 58.5%; Score 24; DB 22; Length 1001;  
Best Local Similarity 75.0%; Pred. No. 5.8;  
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
2 CAGTTGCTGAGAGCCTGAATACAGGAGCTGCTGAGAC 41  
908 CAGTACTGGAAGGCTGACATGTTAGGAGTCTGTGAGCC 869  
'SUIT 14  
'H94457  
AAH94457 standard; cDNA; 1311 BP.  
AAH94457;  
05-OCT-2001 (first entry)  
Human foetal cDNA, SEQ ID NO: 1144.

XX Human; foetal protein; cytostatic; immunosuppressive; immunostimulant;  
KW isotropic; neuroprotective; thrombolytic; osteopathic; antiinflammatory;  
KW gene therapy; antisense therapy; cancer; immune disorder;  
KW growth disorder; osteoporosis; thrombolytic disorder;  
KW nervous system disorder; inflammation; expressed sequence tag; EST; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO200155339-A2.  
XX  
XX 02-AUG-2001.  
XX  
XX 25-JAN-2001; 2001WO-US02723.  
XX  
XX 25-JAN-2000; 2000US-0491404.  
PR 15-SEP-2000; 2000US-0663870.  
PR 06-NOV-2000; 2000US-0707351.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Yeung G, Ford JE, Boyle BJ, Arterburn MC, Drmanac RA, Tang YT;  
PI Liu C, Asundi V, Zhou P, Werhman T;  
PI  
XX WPI; 2001-465571/50.  
DR P-PSDB; AAM06782.  
DR  
XX Novel fetal proteins useful for the treatment and diagnosis of diseases  
PT associated with dysfunction of the protein e.g. cancers, immune  
PT disorders, growth disorders, thrombolytic disorders, nervous system  
PT disorders and inflammation -  
XX  
XX Example 3; Page 642; 715pp; English.  
XX  
XX The invention relates to novel foetal polypeptides encoded by  
CC polynucleotides comprising one of 477 sequences fully defined in the  
CC specification. The foetal polynucleotides and polypeptides are  
CC useful in the treatment and diagnosis of diseases such as cancers,  
CC immune disorders, growth disorders (e.g. osteoporosis), thrombolytic  
CC disorders, nervous system disorders and inflammation. The present  
CC sequence is a full length cDNA which was assembled using expressed  
CC sequence tags (ESTs) found to be expressed in human foetal tissue  
CC cDNA libraries as seeds.  
XX  
SQ Sequence 1311 BP; 375 A; 254 C; 213 G; 469 T; 0 other;  
Query Match 58.5%; Score 24; DB 22; Length 1311;  
Best Local Similarity 75.0%; Pred. No. 6.2;  
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
Qy 2 CAGTTGCTGAGAGCCTGAATACAGGAGCTGCTGAGAC 41  
Db 1223 CAGTACTGGAAGGCTGACATGTTAGGAGTCTGTGAGCC 1262  
RESULT 15  
AAQ84051  
ID AAQ84051 standard; DNA; 3565 BP.  
XX  
XX AC AAQ84051;  
XX  
XX 26-OCT-1995 (first entry)  
XX  
XX Sequence encoding melanoma inhibiting protein (human).  
XX  
XX Melanoma; inhibition; cancer; melanoma; glioblastoma; neuroblastoma;  
KW small cell lung cancer; neuroectodermal tumours; immunosuppressant;  
KW phytohaemagglutinin; lymphocyte; interleukin 2; IL-2; detection;  
XX probe; fusion protein; ss.  
XX  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers

March completed: July 22, 2003, 12:24:12  
 Job time : 175.5 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

|| nucleic - nucleic search, using sw model

n on: July 22, 2003, 11:16:51 ; Search time 36 seconds  
(without alignments)  
349,271 Million cell updates/sec

tlc: U95626-A-AT-42723

rfect score: 41

quence: 1 tcaattgctgagaagctga.....atacaggactgcctgagac 41

oring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

arched: 441362 seqs, 153338381 residues

tal number of hits satisfying chosen parameters: 882724

nimum DB seq length: 0

ximum DB seq length: 2000000000

st-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

tabase : Issued patents NA:\*

1: /cgn2\_6/ptodata/1/ina/5A COMB.seq:\*

2: /cgn2\_6/ptodata/1/ina/5B COMB.seq:\*

3: /cgn2\_6/ptodata/1/ina/6A COMB.seq:\*

4: /cgn2\_6/ptodata/1/ina/6B COMB.seq:\*

5: /cgn2\_6/ptodata/1/ina/PCUTS COMB.seq:\*

6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

sult No.	Score	Query Match	Length	DB	ID	Description
1	24	58.5	3565	1	US-08-578-649-3	Sequence 3, Appli
2	22.4	54.6	5892	3	US-08-755-587-27	Sequence 27, Appl
3	22.4	54.6	8353	3	US-08-611-587-1	Sequence 1, Appli
4	22.4	54.6	15328	2	US-08-888-497-33	Sequence 33, Appl
5	22.4	54.6	15328	4	US-08-382-230-33	Sequence 33, Appl
6	22.4	54.6	15328	5	PCT-US94-07926-33	Sequence 33, Appl
7	22.4	54.6	87350	4	US-08-781-891-79	Sequence 79, Appl
8	22.4	54.6	87543	4	US-09-791-211-3	Sequence 3, Appli
9	21.8	53.2	3224	2	US-08-965-729A-2	Sequence 2, Appli
10	21.8	53.2	3507	2	US-08-775-009-36	Sequence 36, Appl
11	21.4	52.2	43950	4	US-08-735-934A-3	Sequence 3, Appli
12	21.2	51.7	12141	4	US-08-488-671-10	Sequence 10, Appl
13	21	51.2	1927	4	US-08-952-365-3	Sequence 3, Appli
14	21	51.2	2684	1	US-08-367-227-1	Sequence 5, Appli
15	21	51.2	2686	4	US-08-952-365-5	Sequence 5, Appli
16	20.8	50.7	75	4	US-09-461-697-151	Sequence 151, App
17	20.8	50.7	542	4	US-09-461-697-136	Sequence 136, App
18	20.8	50.7	1320	1	US-08-599-352-84	Sequence 84, Appl
19	20.8	50.7	1320	1	US-08-436-074-57	Sequence 57, Appl
20	20.8	50.7	1320	5	PCT-US96-06352-84	Sequence 84, Appl
21	20.8	50.7	1320	5	PCT-US96-06583-84	Sequence 84, Appl
22	20.8	50.7	1400	4	US-09-245-281-40	Sequence 40, Appl
23	20.8	50.7	1400	4	US-09-207-359B-40	Sequence 40, Appl
24	20.8	50.7	5162	2	US-08-916-917-13	Sequence 13, Appl
25	20.8	50.7	5162	3	US-09-225-170-13	Sequence 13, Appl
26	20.8	50.7	8133	4	US-09-659-791A-10	Sequence 10, Appl
27	20.8	50.7	32042	4	US-09-245-281-44	Sequence 44, Appl

28	20.8	50.7	36651	4	US-09-738-894A-3	Sequence 3, Appli
29	20.8	50.7	62804	4	US-09-800-960-3	Sequence 3, Appli
30	20.8	50.7	162450	4	US-09-345-882-1	Sequence 1, Appli
31	20.8	50.7	176373	3	US-09-138-155-17	Sequence 17, Appl
32	20.6	50.2	3441	4	US-09-026-033-17	Sequence 17, Appl
33	20.4	49.8	45716	4	US-08-985-048-5	Sequence 5, Appli
34	20.4	49.8	45789	4	US-08-985-048-6	Sequence 6, Appli
35	20.2	49.3	393	4	US-09-574-141A-97	Sequence 97, Appl
36	20.2	49.3	399	4	US-09-574-141A-94	Sequence 94, Appl
37	20.2	49.3	848	4	US-08-905-223-27	Sequence 27, Appl
38	20.2	49.3	848	4	US-09-247-155-27	Sequence 27, Appl
39	20.2	49.3	862	4	US-09-289-349-6	Sequence 6, Appli
C 40	20.2	49.3	907	3	US-09-081-320-39	Sequence 39, Appl
C 41	20.2	49.3	907	4	US-09-574-141A-39	Sequence 39, Appl
C 42	20.2	49.3	907	4	US-09-707-780-39	Sequence 39, Appl
43	20.2	49.3	6485	3	US-09-081-320-2	Sequence 2, Appli
44	20.2	49.3	6485	4	US-09-574-141A-2	Sequence 2, Appli
45	20.2	49.3	6485	4	US-09-707-780-2	Sequence 2, Appli

#### ALIGNMENTS

RESULT 1  
US-08-578-649-3  
; Sequence 3, Application US/08578649  
; Patent No. 5770366  
; GENERAL INFORMATION:  
; APPLICANT: Ulrich Bogdan  
; APPLICANT: Reinhard Buttner  
; APPLICANT: Brigitte Kaluza  
; TITLE OF INVENTION: MELANOMA-INHIBITING PROTEIN  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Felfe & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/578,649  
; FILING DATE: 29-July-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE P 43 24 247.2  
; FILING DATE: 20-July-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Andrew L. Tiajoloif  
; REGISTRATION NUMBER: 31,575  
; REFERENCE/DOCKET NUMBER: BOER 1035-PFF/ALT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 688-9200  
; TELEFAX: (212) 838-3884  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3565 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: sig peptide  
; LOCATION: 1378..1449  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 1378..1504  
; FEATURE:

NAME/KEY: exon  
LOCATION: 1586..1719  
FEATURE:  
NAME/KEY: exon  
LOCATION: 2804..2914  
FEATURE:  
NAME/KEY: exon  
LOCATION: 3232..3252  
FEATURE:  
NAME/KEY: -  
LOCATION: one-of(2216)  
OTHER INFORMATION: /note= "N in position 2216  
OTHER INFORMATION: denotes an indefinite number ans sequence of  
OTHER INFORMATION: nucleotides"  
3-08-578-649-3

Query Match 58.5%; Score 24; DB 1; Length 3565;  
Best Local Similarity 75.0%; Pred. No. 0.78;  
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

2 CAGTTCGAGAGCCTGAAATACCAAGGACTGCTGAGAC 41  
|||||  
519 CAGTTCAGAGGCTGAGTGGAGGATTGCTGAGTC 558

RESULT 2  
3-08-755-587-27/c  
Sequence 27, Application US/08755587  
Patent No. 6045997  
GENERAL INFORMATION:  
APPLICANT: Futreal, Phillip A  
APPLICANT: Wooster, Richard F  
APPLICANT: Ashworth, Alan  
APPLICANT: Stratton, Michael R  
TITLE OF INVENTION: Materials and methods relating to the  
TITLE OF INVENTION: identification and sequencing of the BRCA2 cancer  
TITLE OF INVENTION: susceptibility gene and uses thereof.  
NUMBER OF SEQUENCES: 222  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Bell Seltzer Park & Gibson  
STREET: 310 UCB Plaza, 3605 Glenwood Avenue, PO Drawer 31107  
CITY: Raleigh  
STATE: NC  
COUNTRY: USA  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/755,587  
FILING DATE: 25-NOV-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9523959.6  
FILING DATE: 23-NOV-1995  
APPLICATION DATA:  
APPLICATION NUMBER: GB 9525555.0  
FILING DATE: 14-DEC-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9617961.9  
FILING DATE: 28-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Kenneth D Sibley  
REGISTRATION NUMBER: 31,665  
REFERENCE/DOCKET NUMBER: 5405-135  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5892 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:

NAME/KEY: CDS  
LOCATION: 483..5412  
FEATURE:  
NAME/KEY: exon  
LOCATION: 481..5412  
US-08-755-587-27

Query Match 54.6%; Score 22.4; DB 3; Length 5892;  
Best Local Similarity 72.5%; Pred. No. 4.1;  
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 CAGTTCGAGAGCCTGAAATACCAAGGACTGCTGAGAC 41  
|||||  
DB 231 CAGTTCAGAGGCTGAGTGGAGGATTGCTGAGCC 192

RESULT 3  
US-08-611-587-1  
Sequence 1, Application US/08611587  
Patent No. 6150091  
GENERAL INFORMATION:  
APPLICANT: PANDOLFO, MASSIMO  
APPLICANT: MONTERMINI, LAURA  
APPLICANT: MOLITO, MARIA D.  
APPLICANT: Koenig, Michael  
APPLICANT: Campuzano, Victoria  
APPLICANT: Cossee, Mireille  
TITLE OF INVENTION: Direct Diagnosis of Friedreich Ataxia  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fulbright & Jaworski L.L.P. Patent Dept.  
STREET: 1301 McKinney, Suite 5100  
CITY: Houston  
STATE: Texas  
COUNTRY: U.S.  
ZIP: 77010  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/611,587  
FILING DATE: 03-MAR-1996  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: Brashears-Macatee, Sarah J.  
REGISTRATION NUMBER: 38,087  
REFERENCE/DOCKET NUMBER: D-5901  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 713-651-5620  
TELEFAX: 713-651-5246  
TELEX: 76-2829  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8353 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapien  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: 9q13  
UNITS: bp  
US-08-611-587-1

Query Match 54.6%; Score 22.4; DB 3; Length 8353;  
Best Local Similarity 72.5%; Pred. No. 4.6;  
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

1/ 2 CAGTTCCTGAGAGCTGAAATACCAAGGACTGCTTGAGAC 41  
|||||  
1972 CAGCTACTGGAAGGCTGAAATGGAAGGACTGCTTGAGCC 2011

## RESULT 4

US-08-888-497-33

Sequence 33, Application US/08888497

Patent No. 5972677

## GENERAL INFORMATION:

APPLICANT: Tischfield, Jay A.

APPLICANT: Seilhamer, Jeffrey J.

TITLE OF INVENTION: Mammalian Phospholipase A2 Nucleotide

TITLE OF INVENTION: Sequences and Low Molecular Weight Amino Acid Sequences

TITLE OF INVENTION: Encoded Thereby, Antisense Sequences and Nucleotide

TITLE OF INVENTION: Sequences Having Internal Ribosome Binding Sites

NUMBER OF SEQUENCES: 44

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ruden, Barnett, McClosky, Smith, Schuster &amp;

ADDRESSEE: Russell PA

STREET: 200 East Broward Boulevard

CITY: Fort Lauderdale

STATE: FL

COUNTRY: USA

ZIP: 33301

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/888,497

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/651,405

FILING DATE:

APPLICATION NUMBER: US 08/097,354

FILING DATE: 26-JUL-1993

ATTORNEY/AGENT INFORMATION:

NAME: Manso, Peter J.

REGISTRATION NUMBER: 32,264

REFERENCE/DOCKET NUMBER: IN21044-5

TELEPHONE: 305-527-2498

TELEPHONE: 305-764-4996

INFORMATION FOR SEQ ID NO: 33:

SEQUENCE CHARACTERISTICS:

LENGTH: 15328 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-08-888-497-33

Query Match 54.6%; Score 22.4; DB 2; Length 15328;

Best Local Similarity 72.5%; Pred. No. 5.4;

Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

1/ 2 CAGTTCCTGAGAGCTGAAATACCAAGGACTGCTTGAGAC 41

|||||

1972 CAGCTACTGGAAGGCTGAAATGGAAGGACTGCTTGAGCC 2011

## RESULT 5

US-09-362-230-33

Sequence 33, Application US/09362230

Patent No. 6352849

## GENERAL INFORMATION:

APPLICANT: Tischfield, Jay A.

APPLICANT: Seilhamer, Jeffrey J.

TITLE OF INVENTION: Mammalian Phospholipase A2 Nucleotide

TITLE OF INVENTION: Sequences and Low Molecular Weight Amino Acid Sequences

TITLE OF INVENTION: Encoded Thereby, Antisense Sequences and Nucleotide

TITLE OF INVENTION: Sequences Having Internal Ribosome Binding Sites

1/ TITLE OF INVENTION: Encoded Thereby, Antisense Sequences and Nucleotide  
TITLE OF INVENTION: Sequences Having Internal Ribosome Binding Sites  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Ruden, Barnett, McClosky, Smith, Schuster &amp;

ADDRESSEE: Russell PA

STREET: 200 East Broward Boulevard

CITY: Fort Lauderdale

STATE: FL

COUNTRY: USA

ZIP: 33301

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/362,230

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/888,497

FILING DATE:

APPLICATION NUMBER: US 08/097,354

FILING DATE: 26-JUL-1993

ATTORNEY/AGENT INFORMATION:

NAME: Manso, Peter J.

REGISTRATION NUMBER: 32,264

REFERENCE/DOCKET NUMBER: IN21044-5

TELEPHONE: 305-527-2498

TELEPHONE: 305-764-4996

INFORMATION FOR SEQ ID NO: 33:

SEQUENCE CHARACTERISTICS:

LENGTH: 15328 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-09-362-230-33

Query Match 54.6%; Score 22.4; DB 4; Length 15328;

Best Local Similarity 72.5%; Pred. No. 5.4;

Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 CAGTTCCTGAGAGCTGAAATACCAAGGACTGCTTGAGAC 41

|||||

Db 8735 CAGCTACTGGAAGGCTGAAATGGAAGGACTGCTTGAGCC 8774

## RESULT 6

PCT-US94-07926-33

Sequence 33, Application PC/TUS9407926

## GENERAL INFORMATION:

APPLICANT: Tischfield, Jay A.

APPLICANT: Seilhamer, Jeffrey J.

TITLE OF INVENTION: Mammalian Phospholipase A2 Nucleotide

TITLE OF INVENTION: Sequences and Low Molecular Weight Amino Acid Sequences

TITLE OF INVENTION: Encoded Thereby, Antisense Sequences and Nucleotide

TITLE OF INVENTION: Sequences Having Internal Ribosome Binding Sites

NUMBER OF SEQUENCES: 44

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ruden, Barnett, McClosky, Smith, Schuster &amp;

ADDRESSEE: Russell PA

STREET: 200 East Broward Boulevard

CITY: Fort Lauderdale

STATE: FL

COUNTRY: USA

ZIP: 33301

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/07926  
FILING DATE: 15-JUL-1994  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/097,354  
FILING DATE: 26-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Manso, Peter J.  
REGISTRATION NUMBER: 32,264  
REFERENCE/DOCKET NUMBER: IN21044-5  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 305-527-2498  
TELEFAX: 305-764-4996  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15328 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
CT-US94-07926-33

Query Match 54.6%; Score 22.4; DB 5; Length 15328;  
Best Local Similarity 72.5%; Pred. No. 5.4;  
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

1y 2 CAGTTGCTGAGAGCTGGAATACCAAGGACTGCTGAGAC 41  
|||||  
1b 8735 CAGTACTCAGAGGCTGAGATAGGAGGATCAGCTGAGCC 8774

## RESULT 7

US-08-781-891-79/c  
Sequence 79, Application US/08781891  
Patent No. 6090620

GENERAL INFORMATION:  
APPLICANT: Fu, Ying-Rui  
APPLICANT: Yu, Chang-En  
APPLICANT: Oshima, Junko  
APPLICANT: Mulligan, John T.  
APPLICANT: Schellenberg, Gerald D.  
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO  
TITLE OF INVENTION: WERNER'S SYNDROME  
NUMBER OF SEQUENCES: 209  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/781,891  
FILING DATE: 27-DEC-1996  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 6090620tenburg Ph.D., Carol  
REGISTRATION NUMBER: 39,317  
REFERENCE/DOCKET NUMBER: 240052.419  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 79:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 87350 base pairs  
TYPE: nucleic acid

; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-781-891-79

Query Match 54.6%; Score 22.4; DB 3; Length 87350;  
Best Local Similarity 72.5%; Pred. No. 8.9; Indels 0; Gaps 0;  
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 CAGTTGCTGAGAGCTGGAATACCAAGGACTGCTGAGAC 41  
|||||  
Db 63492 CAGTACTCAGGATCTGAGATAAGAGGACTGCTTGAGCC 63453

## RESULT 8

US-09-791-211-3/c  
Sequence 3, Application US/09791211  
Patent No. 6448080  
GENERAL INFORMATION:  
APPLICANT: Donna T. Ward  
APPLICANT: Andrew T. Watt  
TITLE OF INVENTION: ANTISENSE MODULATION OF WEN EXPRESSION  
FILE REFERENCE: RTS-0205  
CURRENT APPLICATION NUMBER: US/09/791,211  
CURRENT FILING DATE: 2001-02-23  
NUMBER OF SEQ ID NOS: 90  
SEQ ID NO 3  
LENGTH: 87543

; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: 7421  
; OTHER INFORMATION: unknown  
; NAME/KEY: unsure  
; LOCATION: 7427  
; OTHER INFORMATION: unknown  
; NAME/KEY: unsure  
; LOCATION: 11609  
; OTHER INFORMATION: unknown  
; NAME/KEY: unsure  
; LOCATION: 12605  
; OTHER INFORMATION: unknown  
; NAME/KEY: unsure  
; LOCATION: 12742  
; OTHER INFORMATION: unknown  
; NAME/KEY: unsure  
; LOCATION: 29370  
; OTHER INFORMATION: unknown  
; NAME/KEY: unsure  
; LOCATION: 29422  
; OTHER INFORMATION: unknown  
; NAME/KEY: unsure  
; LOCATION: 29979  
; OTHER INFORMATION: unknown  
; NAME/KEY: unsure  
; LOCATION: 29980  
; OTHER INFORMATION: unknown  
; NAME/KEY: unsure  
; LOCATION: 29981  
; OTHER INFORMATION: unknown  
; NAME/KEY: unsure  
; LOCATION: 30136  
; OTHER INFORMATION: unknown  
; NAME/KEY: unsure  
; LOCATION: 30140  
; OTHER INFORMATION: unknown  
; NAME/KEY: unsure  
; LOCATION: 31205  
; OTHER INFORMATION: unknown  
; NAME/KEY: unsure  
; LOCATION: 31206  
; OTHER INFORMATION: unknown  
; NAME/KEY: unsure

LOCATION: 31592  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 33095  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 33160  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 34066  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 34072  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 36816  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 39020  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 42164  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 42459  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 46808  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 46823  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 46826  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 47291  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 52786  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 52787  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 53384  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 54684  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 59215  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 59235  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 59242  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 63290  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 66614  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 68660  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 68697  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 68718

OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 68733  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 68739  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 69785  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 79134  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 79198  
OTHER INFORMATION: unknown  
NAME/KEY: unsure  
LOCATION: 86336  
OTHER INFORMATION: unknown  
OTHER INFORMATION:  
US-09-791-211-3

Query Match 54.6%; Score 22.4; DB 4; Length 87543;  
Best Local Similarity 72.5%; Pred. No. 8.9;  
Matches 29; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 CAGTGTCTGAGAGCTGGAATACACGAGCTGCTGAGAC 41  
|||||  
Db 63685 CAGCTACTCAGGACTGTGAGATAAGAGGACTGCTGAGCC 63646  
|||||

RESULT 9  
US-08-965-729A-2/c  
; Sequence 2, Application US/08965729A  
; Patent No. 6200751  
; GENERAL INFORMATION:  
; APPLICANT: Jian-Ming Gu and Charles T. Esmon  
; TITLE OF INVENTION: ENDOTHELIUM SPECIFIC EXPRESSION  
; TITLE OF INVENTION: REGULATED BY EPCR CONTROL ELEMENTS  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Patrea L. Pabst  
; STREET: 2800 One Atlantic Center, 1201 West Peachtree Street  
; CITY: Atlanta  
; STATE: GA  
; COUNTRY: USA  
; ZIP: 30309-4530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/965,729A  
; FILING DATE: 07-NOV-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/030,718  
; FILING DATE: 08-NOV-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pabst, Patrea L.  
; REGISTRATION NUMBER: 31,284  
; REFERENCE/DOCKET NUMBER: OMRF 164 PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 404-873-8794  
; TELEFAX: 404-873-8795  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3224 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)



HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: /note= "Nucleotides 2272 through 2729 are a  
OTHER INFORMATION: thrombin responsive element"; Human  
US-08-965-729A-2

Query Match 53.2%; Score 21.8; DB 4; Length 3224;  
Best Local Similarity 70.7%; Pred. No. 6.2; Mismatches 12; Indels 0; Gaps 0;  
Matches 29; Conservative 0;

1 TCAGTGTCTGAGAGCCTGAAATACAGGACTGCTGAGAC 41  
|||||  
198 TCAGTACTCAGAGGCTGAGTAGAGGATGCTTGAGCC 158  
|||||

RESULT 10  
US-08-775-009-36/c  
Sequence 36, Application US/08775009  
Patent No. 5935783  
GENERAL INFORMATION:  
APPLICANT: Gong, Weiliang  
APPLICANT: Emanuel, Beverly S.  
APPLICANT: Budarf, Marcia L.  
APPLICANT: Roe, Bruce  
TITLE OF INVENTION: No. 5935783el Genes Mapping in the Digeorge and  
TITLE OF INVENTION: Veliocardiofacial Syndrome Minimal Critical Region  
NUMBER OF SEQUENCES: 45  
CORRESPONDENCE ADDRESS:  
ADDRESS: Woodcock Washburn Kurtz Mackiewicz &  
ADDRESS: No. 5935783ris, LLP  
STREET: One Liberty Place - 46th Floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: U.S.A.  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/775,009  
FILING DATE: 27-DEC-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Trujillo, Doreen Yanko  
REGISTRATION NUMBER: 35,719  
REFERENCE/DOCKET NUMBER: CH-0681  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-3100  
TELEFAX: (215) 568-3439  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3507 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: exon  
LOCATION: 163..423  
FEATURE:  
NAME/KEY: exon  
LOCATION: 532..783  
FEATURE:  
NAME/KEY: exon  
LOCATION: 1351..1458  
FEATURE:  
NAME/KEY: CDS  
LOCATION: join(163..423, 532..783, 1351..1458)  
US-08-775-009-36

Query Match 53.2%; Score 21.8; DB 2; Length 3507;  
Best Local Similarity 78.8%; Pred. No. 6.3; Mismatches 7; Indels 0; Gaps 0;  
Matches 26; Conservative 0;

2 CAGTGTCTGAGAGCCTGAAATACAGGACTGC 34  
|||||  
1940 CAGGGCTGGAGAGCTGCAACGCGGAGACTGC 1908  
|||||

RESULT 11  
US-09-735-934A-3  
Sequence 3, Application US/09735934A  
Patent No. 6372468  
GENERAL INFORMATION:  
APPLICANT: Li, Jiajin et al  
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
TITLE OF INVENTION: THEREOF  
FILE REFERENCE: CL000851  
CURRENT APPLICATION NUMBER: US/09/735,934A  
CURRENT FILING DATE: 2000-12-14  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 43950  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-735-934A-3

Query Match 52.2%; Score 21.4; DB 4; Length 43950;  
Best Local Similarity 71.8%; Pred. No. 19; Mismatches 11; Indels 0; Gaps 0;  
Matches 28; Conservative 0;

3 AGTGTCTGAGAGCCTGAAATACAGGACTGCTGAGAC 41  
|||||  
36201 AGCTGCTTAGAGGCTGAGATGGAGGATGCTTGAGCC 36239  
|||||

RESULT 12  
US-09-488-671-10/c  
Sequence 10, Application US/09488671A  
Patent No. 6187545  
GENERAL INFORMATION:  
APPLICANT: Robert McKay  
APPLICANT: Madeline M. Butler  
APPLICANT: Jacqueline Wyatt  
APPLICANT: Lex M. Cowsett  
TITLE OF INVENTION: ANTISENSE MODULATION OF PEPCK-CYTOSOLIC EXPRESSION  
FILE REFERENCE: RTS-0123  
CURRENT APPLICATION NUMBER: US/09/488,671A  
CURRENT FILING DATE: 2000-01-21  
NUMBER OF SEQ ID NOS: 177  
SEQ ID NO 10  
LENGTH: 12141  
TYPE: DNA  
ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (5895)...(6118)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (6440)...(6621)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (7206)...(7409)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (7682)...(7869)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (8444)...(8606)  
FEATURE:

NAME/KEY: CDS  
LOCATION: (9418)...(9642)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (9735)...(9866)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (10502)...(10597)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (10701)...(11155)  
IS-09-488-671-10

Query Match 51.7%; Score 21.2; DB 4; Length 12141;  
Best Local Similarity 76.5%; Pred. No. 16;  
Matches 26; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

'Y 7 GCTGAGAAGCCTGAATACACAGACTGCTGAGA 40  
|||||  
'b 6805 GCTGACAGGCTGAGCCACAGACTTGACAGA 6772  
|||||

RESULT 13  
S-08-952-365-3  
Sequence 3, Application US/08952365  
Patent No. 6274311  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: A Method and Nucleotide Sequence for  
Patent No. 6274311  
TITLE OF INVENTION: Transforming Microorganisms  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BERESKIN & PARR  
STREET: 40 King Street West  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5H 3Y2  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/952.365  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Kurdydk, Linda M.  
REGISTRATION NUMBER: 34,971  
REFERENCE/DOCKET NUMBER: 6580-63  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 364-7311  
TELEFAX: (416) 361-1398  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1927 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: Lactococcus lactis  
IMMEDIATE SOURCE:  
CLONE: EML  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 267..1832  
IS-08-952-365-3

Query Match 51.2%; Score 21; DB 4; Length 1927;  
Best Local Similarity 73.0%; Pred. No. 11;

Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2 CAGTTGCTGAGAAGCCTGAATACACAGACTGCTGAGA 38  
|||||  
Db 1735 CAGTTGCTAAGAAGCTCAGAGACAGAGTCTTACTGA 1771  
|||||

RESULT 14  
US-08-367-227-1  
; Sequence 1, Application US/08367227  
; Patent No. 5587304  
; GENERAL INFORMATION:  
; APPLICANT: BARRE, PIERRE  
; APPLICANT: DEQUIN, SYLVIE  
; APPLICANT: ANSANY, VIRGINIE  
; TITLE OF INVENTION: CLONING AND EXPRESSION OF THE GENE FOR  
; TITLE OF INVENTION: THE MALOLACTIC ENZYME OF LACTOCOCCUS LACTIS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT  
; STREET: 1755 S. JEFFERSON DAVIS HWY, SUITE 400  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/367,227  
; FILING DATE: 17-JAN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR PCT/FR94/00589  
; FILING DATE: 18-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: FR 93/06003  
; FILING DATE: 18-MAY-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 384-42-0 PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2684 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; ANTI-SENSE: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 486..2085  
; FEATURE:  
; NAME/KEY: -35\_signal  
; LOCATION: 392..397  
; FEATURE:  
; NAME/KEY: -10\_signal  
; LOCATION: 416..421  
US-08-367-227-1

Query Match 51.2%; Score 21; DB 1; Length 2684;  
Best Local Similarity 73.0%; Pred. No. 13;  
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2 CAGTTGCTGAGAAGCCTGAATACACAGACTGCTGAGA 38  
|||||

0b 1991 CAGTTGCTAAGAAAGCTCAAGAACCAAGGCTTACTGA 2027

RESULT 15  
US-08-952-365-5  
Sequence 5, Application US/08952365  
Patent No. 6274311  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: A Method and Nucleotide Sequence for  
Patent No. 6274311  
TITLE OF INVENTION: Transforming Microorganisms  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BERESKIN & PARR  
STREET: 40 King Street West  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5H 3Y2  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/952,365  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Kurdydyk, Linda M.  
REGISTRATION NUMBER: 34,971  
REFERENCE/DOCKET NUMBER: 6580-63  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 364-7311  
TELEFAX: (416) 361-1398  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2686 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: Lactococcus lactis  
IMMEDIATE SOURCE:  
CLONE: mles  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 467..2089  
US-08-952-365-5

Query Match 51.2%; Score 21; DB 4; Length 2686;  
Best Local Similarity 73.0%; Pred No.13;  
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

2y 2 CAGTTGCTGAGAGCTGAATACCAGGACTGCCTGA 38  
|||||  
0b 1992 CAGTTGCTAAGAAAGCTCAAGAACCAAGGCTTACTGA 2028

Search completed: July 22, 2003, 11:18:24  
Job time : 38 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.  
M nucleic - nucleic search, using sw model  
Run on: July 22, 2003, 11:16:53 ; Search time 130 Seconds  
(without alignments)  
650.639 Million cell updates/sec

File: U95626-A-AT-42723  
Perfect score: 41  
Sequence: 1 tcagtctgtagagcctgga.....ataccaggactgctgagac 41  
Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 1439767 seqs, 1031500376 residues  
Total number of hits satisfying chosen parameters: 2879534  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubna/US09\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/1/pubna/US09\_NEW\_PUB.seq.\*  
11: /cgn2\_6/ptodata/1/pubna/US09\_NEW\_PUB.seq.\*  
12: /cgn2\_6/ptodata/1/pubna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubna/US09\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubna/US10\_NEW\_PUB.seq.\*  
15: /cgn2\_6/ptodata/1/pubna/US10\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubna/US60\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/1/pubna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39.4	96.1	143068	11	US-09-967-768A-316
2	24.6	60.0	610	15	US-10-027-632-224761
3	24	58.5	544	15	US-10-027-632-66000
4	24	58.5	544	15	US-10-027-632-66001
5	24	58.5	544	15	US-10-027-632-66002
6	24	58.5	544	15	US-10-027-632-298997
7	24	58.5	544	15	US-10-027-632-298998
8	24	58.5	544	15	US-10-027-632-298999
9	24	58.5	1206	15	US-10-027-632-210607
10	24	58.5	17335	11	US-09-764-847-1280
11	24	58.5	17335	15	US-10-092-154-1280
12	24	58.5	19882	11	US-09-764-847-1281
13	24	58.5	19882	15	US-10-092-154-1281
14	24	58.5	183337	15	US-10-020-141-5
15	24	58.5	202001	10	US-09-734-674-3
16	24	58.5	202001	15	US-10-274-990-3

17	24	58.5	300000	15	US-10-262-552-33	Sequence 33, Appl
18	23.6	57.6	591	15	US-10-198-846-11772	Sequence 11772, A
19	23.6	57.6	654	15	US-10-027-632-234154	Sequence 234154, A
20	23.4	57.1	590	15	US-10-027-632-216678	Sequence 216678, A
21	23.4	57.1	644	15	US-10-027-632-254451	Sequence 254451, A
22	23.4	57.1	731	15	US-10-027-632-152011	Sequence 152011, A
23	23.4	57.1	2149	15	US-10-027-632-103171	Sequence 103171, A
24	23.4	57.1	2149	15	US-10-027-632-103172	Sequence 103172, A
25	23.4	57.1	38374	11	US-09-880-107-3463	Sequence 3463, Ap
26	23.4	57.1	1503841	10	US-09-795-688-1	Sequence 1, Appl
27	23.4	57.1	1503841	10	US-09-795-686-1	Sequence 1, Appl
28	23.4	57.1	1503841	11	US-09-946-807-1	Sequence 1, Appl
29	23.2	56.6	8835	12	US-09-764-891-5494	Sequence 5494, Ap
30	22.8	55.6	1835	15	US-10-027-632-255911	Sequence 255911, A
31	22.6	55.1	413	15	US-10-027-632-55856	Sequence 55856, A
32	22.6	55.1	413	15	US-10-027-632-301953	Sequence 301953, A
33	22.6	55.1	676	15	US-10-027-632-222706	Sequence 222706, A
34	22.6	55.1	676	15	US-10-027-632-222707	Sequence 222707, A
35	22.6	55.1	879	15	US-10-027-632-2178	Sequence 2178, Ap
36	22.6	55.1	118067	14	US-10-081-327-32	Sequence 32, Appl
37	22.6	55.1	145831	11	US-09-969-708-79	Sequence 79, Appl
38	22.6	55.1	145831	11	US-09-954-456-2116	Sequence 2116, Ap
39	22.4	54.6	211	15	US-10-066-543-2056	Sequence 2056, Ap
40	22.4	54.6	322	11	US-09-764-877-2800	Sequence 2800, Ap
41	22.4	54.6	397	15	US-10-198-846-913	Sequence 913, App
42	22.4	54.6	401	15	US-10-027-632-279579	Sequence 279579, A
43	22.4	54.6	497	15	US-10-027-632-214260	Sequence 214260, A
44	22.4	54.6	507	15	US-10-027-632-206966	Sequence 206966, A
45	22.4	54.6	535	15	US-10-198-846-7806	Sequence 7806, Ap

ALIGNMENTS

RESULT 1

US-09-967-768A-316  
; Sequence 316, Application US/09967768A  
; Patent No. US20020150877A1  
; GENERAL INFORMATION:  
; APPLICANT: Augustus, Meena  
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatur  
; FILE REFERENCE: 689290-72  
; CURRENT APPLICATION NUMBER: US/09/967,768A  
; CURRENT FILING DATE: 2001-09-28  
; PRIOR APPLICATION NUMBER: US/60/236,109  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US/60/236,034  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US/60/236,111  
; PRIOR FILING DATE: 2000-09-28  
; NUMBER OF SEQ ID NOS: 325  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 316  
; LENGTH: 143068  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-967-768A-316

Query Match 96.1%; Score 39.4; DB 11; Length 143068;  
Best Local Similarity 97.6%; Pred. No. 4.4e-06;  
Matches 40; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCAGTCTGAGAGCCTGGAATACACAGGACTGCTGAGAC 41  
|||||  
Db 42703 TCAGTCTGAGAGCCTGGAATACACAGGACTGCTGAGAC 42743

RESULT 2

US-10-027-632-224761/c  
; Sequence 224761, Application US/10027632  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.

;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
;; TITLE OF INVENTION: Polymorphisms in the Human Genome

;; FILE REFERENCE: 108827.129  
;; CURRENT FILING DATE: 2002-04-30

;; PRIOR APPLICATION NUMBER: US 60/218,006  
;; PRIOR FILING DATE: 2000-07-12

;; PRIOR APPLICATION NUMBER: US 60/198,676  
;; PRIOR FILING DATE: 2000-04-20

;; PRIOR APPLICATION NUMBER: US 60/193,483  
;; PRIOR FILING DATE: 2000-03-29

;; PRIOR APPLICATION NUMBER: US 60/185,218  
;; PRIOR FILING DATE: 2000-02-24

;; PRIOR APPLICATION NUMBER: US 60/167,363  
;; PRIOR FILING DATE: 1999-11-23

;; PRIOR APPLICATION NUMBER: US 60/156,358  
;; PRIOR FILING DATE: 1999-09-28

;; PRIOR APPLICATION NUMBER: US 60/146,002  
;; PRIOR FILING DATE: 1999-08-09

;; NUMBER OF SEQ ID NOS: 325720  
;; SOFTWARE: FastSeq for Windows Version 4.0

;; SEQ ID NO 224761  
;; LENGTH: 610

;; TYPE: DNA  
;; ORGANISM: Human

;; US-10-027-632-66001

Query Match 60.8%; Score 24.6; DB 15; Length 610;  
Best Local Similarity 76.9%; Pred. No. 1.9;

Matches 30; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

1 TCAGTTGCTGAGAGCTGAAATACAGGACTGCTTGAG 39

80 TCAGTACTAAAGGCTGAAATACAGGACTGCTTGAG 42

RESULT 3

US-10-027-632-66000  
;; Sequence 66000, Application US/10027632

;; GENERAL INFORMATION:  
;; APPLICANT: Wang, David G.

;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
;; TITLE OF INVENTION: Polymorphisms in the Human Genome

;; FILE REFERENCE: 108827.129  
;; CURRENT APPLICATION NUMBER: US/10/027,632

;; CURRENT FILING DATE: 2002-04-30  
;; PRIOR APPLICATION NUMBER: US 60/218,006

;; PRIOR FILING DATE: 2000-07-12  
;; PRIOR APPLICATION NUMBER: US 60/198,676

;; PRIOR FILING DATE: 2000-04-20  
;; PRIOR APPLICATION NUMBER: US 60/193,483

;; PRIOR FILING DATE: 2000-03-29  
;; PRIOR APPLICATION NUMBER: US 60/185,218

;; PRIOR FILING DATE: 2000-02-24  
;; PRIOR APPLICATION NUMBER: US 60/167,363

;; PRIOR FILING DATE: 1999-11-23  
;; PRIOR APPLICATION NUMBER: US 60/156,358

;; PRIOR FILING DATE: 1999-09-28  
;; PRIOR APPLICATION NUMBER: US 60/146,002

;; PRIOR FILING DATE: 1999-08-09  
;; NUMBER OF SEQ ID NOS: 325720

;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 66000

;; LENGTH: 544  
;; TYPE: DNA

;; ORGANISM: Human  
;; US-10-027-632-66000

Query Match 58.5%; Score 24; DB 15; Length 544;

Best Local Similarity 75.0%; Pred. No. 3.3;

Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

2 CAGTTGCTGAGAGCTGAAATACAGGACTGCTTGAG 41

Db 127 CAGTACTCAGAGGCTGAGTGCAGGATTGCTTGAGAC 166

RESULT 4

US-10-027-632-66001

;; Sequence 66001, Application US/10027632

;; GENERAL INFORMATION:  
;; APPLICANT: Wang, David G.

;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
;; TITLE OF INVENTION: Polymorphisms in the Human Genome

;; FILE REFERENCE: 108827.129  
;; CURRENT APPLICATION NUMBER: US/10/027,632

;; CURRENT FILING DATE: 2002-04-30  
;; PRIOR APPLICATION NUMBER: US 60/218,006

;; PRIOR FILING DATE: 2000-07-12  
;; PRIOR APPLICATION NUMBER: US 60/198,676

;; PRIOR FILING DATE: 2000-04-20  
;; PRIOR APPLICATION NUMBER: US 60/193,483

;; PRIOR FILING DATE: 2000-03-29  
;; PRIOR APPLICATION NUMBER: US 60/185,218

;; PRIOR FILING DATE: 2000-02-24  
;; PRIOR APPLICATION NUMBER: US 60/167,363

;; PRIOR FILING DATE: 1999-11-23  
;; PRIOR APPLICATION NUMBER: US 60/156,358

;; PRIOR FILING DATE: 1999-09-28  
;; PRIOR APPLICATION NUMBER: US 60/146,002

;; PRIOR FILING DATE: 1999-08-09  
;; NUMBER OF SEQ ID NOS: 325720

;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 66001

;; LENGTH: 544  
;; TYPE: DNA

;; ORGANISM: Human  
;; US-10-027-632-66001

Query Match 58.5%; Score 24; DB 15; Length 544;

Best Local Similarity 75.0%; Pred. No. 3.3;

Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 2 CAGTTGCTGAGAGCTGAAATACAGGACTGCTTGAGAC 41

Db 127 CAGTACTCAGAGGCTGAGTGCAGGATTGCTTGAGAC 166

RESULT 5

US-10-027-632-66002

;; Sequence 66002, Application US/10027632

;; GENERAL INFORMATION:  
;; APPLICANT: Wang, David G.

;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
;; TITLE OF INVENTION: Polymorphisms in the Human Genome

;; FILE REFERENCE: 108827.129  
;; CURRENT APPLICATION NUMBER: US/10/027,632

;; CURRENT FILING DATE: 2002-04-30  
;; PRIOR APPLICATION NUMBER: US 60/218,006

;; PRIOR FILING DATE: 2000-07-12  
;; PRIOR APPLICATION NUMBER: US 60/198,676

;; PRIOR FILING DATE: 2000-04-20  
;; PRIOR APPLICATION NUMBER: US 60/193,483

;; PRIOR FILING DATE: 2000-03-29  
;; PRIOR APPLICATION NUMBER: US 60/185,218

;; PRIOR FILING DATE: 2000-02-24  
;; PRIOR APPLICATION NUMBER: US 60/167,363

;; PRIOR FILING DATE: 1999-11-23  
;; PRIOR APPLICATION NUMBER: US 60/156,358

;; PRIOR FILING DATE: 1999-09-28  
;; PRIOR APPLICATION NUMBER: US 60/146,002

;; PRIOR FILING DATE: 1999-08-09  
;; NUMBER OF SEQ ID NOS: 325720

;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 66002

;; LENGTH: 544  
;; TYPE: DNA

;; ORGANISM: Human  
;; US-10-027-632-66002



```

; LENGTH: 17335
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-092-154-1280

Query Match      58.5%; Score 24; DB 15; Length 17335;
Best Local Similarity 75.0%; Pred. No. 5.8;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY      2  CAGTGTCTGAGAGGCTGAATATACCGAGGACTGCCTTGAGAC 41
      |||
Db      5266  CAGCTACTCAGAAGGCTGAGACAGGAGGACTGCTTGAGCC 5227

RESULT 12
US-09-764-847-1281/c
; Sequence 1281, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1281
; LENGTH: 19882
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-847-1281

Query Match      58.5%; Score 24; DB 11; Length 19882;
Best Local Similarity 75.0%; Pred. No. 6;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY      2  CAGTGTCTGAGAGGCTGAATATACCGAGGACTGCCTTGAGAC 41
      |||
Db      5270  CAGCTACTCAGAAGGCTGAGACAGGAGGACTGCTTGAGCC 5231

RESULT 13
US-10-092-154-1281/c
; Sequence 1281, Application US/10092154
; Publication No. US20030054375A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009C1
; CURRENT APPLICATION NUMBER: US/10/092,154
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2003
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1281
; LENGTH: 19882
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-092-154-1281

Query Match      58.5%; Score 24; DB 15; Length 19882;
Best Local Similarity 75.0%; Pred. No. 6;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY      2  CAGTGTCTGAGAGGCTGAATATACCGAGGACTGCCTTGAGAC 41
      |||
Db      5270  CAGCTACTCAGAAGGCTGAGACAGGAGGACTGCTTGAGCC 5231

RESULT 14
US-10-020-141-5/c
; Sequence 5, Application US/10020141

```







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

M nucleic - nucleic search, using sw model

Run on: July 22, 2003, 11:16:52 ; Search time 1307 Seconds  
(without alignments)  
508.045 Million cell updates/sec

itle: U95626-A-AT-42723  
erfect score: 41  
equence: 1 tcagttgctgagagcctga.....ataccagagtcgctgagac 41

oring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

earched: 16154066 seqs, 8097743376 residues

otal number of hits satisfying chosen parameters: 32308132

inimum DB seq length: 0  
aximum DB seq length: 2000000000

ost-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

atabase : EST.\*  
1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estnu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_htc.\*  
9: gb\_estl.\*  
10: gb\_est2.\*  
11: gb\_htc.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: gb\_gss.\*  
18: em\_gss\_hum.\*  
19: em\_gss\_inv.\*  
20: em\_gss\_pln.\*  
21: em\_gss\_vrt.\*  
22: em\_gss\_fun.\*  
23: em\_gss\_mus.\*  
24: em\_gss\_nam.\*  
25: em\_gss\_oth.\*  
26: em\_gss\_pro.\*  
27: em\_gss\_rtd.\*

result No.	Score	Match	Length	DB	ID	Description
1	26.6	64.9	352	17	B85322	B85322 RPC111-16C1
2	26.6	64.9	368	17	AQ132663	AQ132663 HS 3051_B
3	25.6	62.4	412	17	AQ181316	AQ181316 HS 3223_A
4	25.6	62.4	471	14	BQ300489	BQ300489 PMO-KT004
5	25.6	62.4	509	17	AQ898508	AQ898508 HS 3127_B
6	25.6	62.4	538	10	AW653656	AW653656 102451_MA

# SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

7	25.6	62.4	699	10	AW978350	AW978350 EST390459
8	25.2	61.5	453	9	AA806631	AA806631 OB71108_S
9	25	61.0	326	10	AW900757	AW900757 CM1-WN100
10	25	61.0	457	10	AW070703	AW070703 Xa29506.X
11	24.6	60.0	836	9	AL575170	AL575170 AL575170
12	24	58.5	315	12	BF935457	BF935457 CM1-NT027
13	24	58.5	337	9	AA369692	AA369692 EST81420
14	24	58.5	374	10	AW386852	AW386852 CM0-PT004
15	24	58.5	384	12	BF930363	BF930363 CM0-GN016
16	24	58.5	384	12	BF995389	BF995389 CM0-GN016
17	24	58.5	384	12	BG002012	BG002012 CM0-GN016
18	24	58.5	418	10	BE351027	BE351027 ht6409.X
19	24	58.5	423	12	BF760350	BF760350 CM1-CT060
20	24	58.5	430	13	BI043650	BI043650 PM4-OT020
21	24	58.5	457	17	AQ423186	AQ423186 CITBI-RI-
22	24	58.5	458	10	AW848768	AW848768 IL3-CT021
23	24	58.5	470	12	BG002336	BG002336 QV4-GN025
24	24	58.5	493	14	BQ300457	BQ300457 PMO-KT004
25	24	58.5	499	10	AW575409	AW575409 UI-HP-BL0
26	24	58.5	512	12	BF858648	BF858648 RC5-PT019
27	24	58.5	524	17	AQ784899	AQ784899 HS 3251_A
28	24	58.5	536	13	BM272482	BM272482 IG4d11.Y
29	24	58.5	560	17	AQ419838	AQ419838 RPCI-11-1
30	24	58.5	571	17	AQ504035	AQ504035 RPCI-11-2
31	24	58.5	586	9	AL708735	AL708735 DXFZP686P
32	24	58.5	636	17	AQ616424	AQ616424 HS 5149_A
33	24	58.5	646	17	AG040016	AG040016 Pan trogl
34	24	58.5	648	17	AG147167	AG147167 Pan trogl
35	24	58.5	663	17	AG149846	AG149846 Pan trogl
36	24	58.5	685	17	AQ075068	AQ075068 CIT-HSP-2
37	24	58.5	826	17	AQ749940	AQ749940 HS 5576_A
38	24	58.5	840	17	AQ898416	AQ898416 HS 3122_B
39	24	58.5	841	14	BQ719041	BQ719041 AGENCOURT
40	23.8	58.0	1047	14	BQ049246	BQ049246 AGENCOURT
41	23.6	57.6	184	12	BF957211	BF957211 PM1-WN120
42	23.6	57.6	370	12	BG197441	BG197441 RST16686
43	23.6	57.6	409	17	B50624	B50624 CIT-HSP-523
44	23.6	57.6	417	17	AQ802266	AQ802266 HS 3136_B
45	23.6	57.6	431	9	AI039408	AI039408 ox40102.5

## ALIGNMENTS

RESULT 1  
B85322  
LOCUS  
DEFINITION  
B85322  
DNA sequence.  
B85322  
B85322.1  
GI:2926454  
B85322  
GSS.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (Bases 1 to 352)  
Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.  
Use of BAC End Sequences for Sequence-Ready Map Building (1998)  
Unpublished (1998)  
Other\_GSSs: RPC111-16C1.8.TV  
Contact: Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: mdaamset@igr.org  
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering/>) or from Research Genetics ([inforesgen.com](http://inforesgen.com)). BAC end search page: [http://www.tigr.org/tdb/humgen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html)  
Seq primer: SP6  
Class: BAC ends.

# FEATURES

Location/Qualifiers  
1..352

source  
/organism="Homo sapiens"  
/db\_xref="GDB:7505825"  
/db\_xref="taxon:9606"  
/clone="RPC1-11-16C18"  
/clone\_lib="RPC1-11"  
/sex="Male"  
/cell\_type="Lymphocytes"  
/note="Vector: pBACe3.6; Site\_1: EcoRI; Site\_2: EcoRI; RPC111 Human Male BAC Library"

BASE COUNT 105 a 72 c 85 g 90 t

# ORIGIN

Query Match 64.9%; Score 26.6; DB 17; Length 352;  
Best Local Similarity 78.0%; Pred. No. 14;  
Matches 32; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

1 TCAGTGTCTGGAAGCTGAATACAGGACTGCTGTGAGAC 41  
85 TCAGTGTCTGGAAGCTGAATGAGGAGACTGCTGTGAGCC 125

# RESULT 2

# LOCUS

AQ132663 368 bp DNA linear GSS 23-SEP-1998  
HS 3051 B2 H11 MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3051 Col=22 Row=P, DNA sequence.

# ACCESSION

# VERSION

# KEYWORDS

# SOURCE

# ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 368)  
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.  
Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

# JOURNAL

# MEDLINE

# COMMENT

Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
99380589  
Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Sequence Tagged Connector  
Plate: 3051 row: P column: 22  
Class: BAC ends  
High quality sequence stop: 368.

# FEATURES

source

1..368  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="Plate=3051 Col=22 Row=P"  
/clone\_lib="CIT Approved Human Genomic Sperm Library D"  
/sex="male"  
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in E-Coli DH10B"

# BASE COUNT

# ORIGIN

107 a 75 c 91 g 92 t 3 others  
Query Match 64.9%; Score 26.6; DB 17; Length 368;  
Best Local Similarity 78.0%; Pred. No. 14;

Matches 32; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

1 TCAGTGTCTGGAAGCTGAATACAGGACTGCTGTGAGAC 41  
87 TCAGTGTCTGGAAGCTGAATGAGGAGACTGCTGTGAGCC 127

# RESULT 3

# AQ181316/c

# LOCUS

AQ181316 412 bp DNA linear GSS 23-OCT-1998  
HS 3223 A1 F11 MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3223 Col=21 Row=K, DNA sequence.

# ACCESSION

# VERSION

# KEYWORDS

# SOURCE

# ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 412)  
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.  
Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

# JOURNAL

# MEDLINE

# COMMENT

Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
99380589  
Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Sequence Tagged Connector  
Plate: 3223 row: K column: 21  
Class: BAC ends  
High quality sequence stop: 412.

# FEATURES

source

1..412  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="Plate=3223 Col=21 Row=K"  
/clone\_lib="CIT Approved Human Genomic Sperm Library D"  
/sex="male"  
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in E-Coli DH10B"

BASE COUNT 97 a 95 c 75 g 144 t 1 others

# ORIGIN

Query Match 62.4%; Score 25.6; DB 17; Length 412;  
Best Local Similarity 77.5%; Pred. No. 33;  
Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

2 CAGTGTCTGAGAGCTGAATACAGGACTGCTGTGAGAC 41  
184 CAGCTACTTGGAGGCTGAAGTAGGAGGACTGCTGTGAGCC 145

# RESULT 4

# LOCUS

BQ300489 471 bp mRNA linear EST 16-MAY-2002  
PMO-KT0041-070201-004-B04 KT0041 Homo sapiens cDNA, mRNA sequence.

# ACCESSION

# VERSION

# KEYWORDS

# SOURCE

# ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 471)  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J., and Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

# TITLE

## JOURNAL

## MEDLINE

## COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PMO&t2=PMO-KT0041-070201-004-b04&t3=2001-02-07&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 11

High quality sequence stop: 468.

Location/Qualifiers

# FEATURES

## source

## 1..471

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="KT0041"

/dev\_stage="Adult"

/note="Organ: bladder tumor; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

low stringency conditions."

ASE COUNT 133 a 119 c 115 g 104 t

# RIGIN

Query Match 62.4%; Score 25.6; DB 14; Length 471;

Best Local Similarity 77.5%; Pred. No. 35;

Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Y 2 CAGTTCCTGAGAGCTGAAATACAGACTGCTGAGAC 41

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Clones may be purchased from Research Genetics (info@resgen.com).

BAC end Web Server: http://www.hnsc.washington.edu

Plate: 3127 row: F column: 20

Seq primer: T7

Class: BAC ends

High quality sequence stop: 509.

Location/Qualifiers

1..509

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="Plate=3127 Col=20 Row=F"

/clone\_lib="CIT Approved Human Genomic Sperm Library D"

/sex="male"

/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in B-Coli DH10B"

BASE COUNT 125 a 121 c 103 g 154 t 6 others

# ORIGIN

## Query Match

## Best Local Similarity

## Matches 31; Conservative

## 0; Mismatches 9; Indels 0; Gaps 0;

## QY 2 CAGTTCCTGAGAGCTGAAATACAGACTGCTGAGAC 41

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

## |||||

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Clones may be purchased from Research Genetics (info@resgen.com).

BAC end Web Server: http://www.hnsc.washington.edu

Plate: 3127 row: F column: 20

Seq primer: T7

Class: BAC ends

High quality sequence stop: 509.

Location/Qualifiers

1..509

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="Plate=3127 Col=20 Row=F"

/clone\_lib="CIT Approved Human Genomic Sperm Library D"

/sex="male"

/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in B-Coli DH10B"

BASE COUNT 125 a 121 c 103 g 154 t 6 others

ORIGIN

Query Match 62.4%; Score 25.6; DB 17; Length 509;

Best Local Similarity 77.5%; Pred. No. 36;

Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 2 CAGTTCCTGAGAGCTGAAATACAGACTGCTGAGAC 41

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Clones may be purchased from Research Genetics (info@resgen.com).

BAC end Web Server: http://www.hnsc.washington.edu

Plate: 3127 row: F column: 20

Seq primer: T7

Class: BAC ends

High quality sequence stop: 509.

Location/Qualifiers

1..509

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="Plate=3127 Col=20 Row=F"

/clone\_lib="CIT Approved Human Genomic Sperm Library D"

/sex="male"

/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in B-Coli DH10B"

BASE COUNT 125 a 121 c 103 g 154 t 6 others

Library made from pooled tissue from lymph node, ovary, fat, hypothalamus, and pituitary."

BASE COUNT 150 a 132 c 123 g 133 t  
 ORIGIN  
 Query Match 62.4%; Score 25.6; DB 10; Length 538;  
 Best Local Similarity 77.5%; Pred. No. 36;  
 Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
 PY 1 TCAGTCTCAGAGCGCTGAATACCGAGCTGCTGAGA 40  
 362 TCAGTCTCAGAGCGCTGACACACACAGCTGCTTAA 401

RESULT 7  
 W978350 699 bp mRNA linear EST 02-JUN-2000  
 DEFINITION EST390459 MAGE resequences, MAGP Homo sapiens cDNA, mRNA sequence.  
 ACCESSION AW978350  
 VERSION AW978350.1 GI:8169615  
 KEYWORDS EST.  
 SOURCE human.

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 699)  
 AUTHORS Hegde,P., Qi,R., Abernathy,K., Gaspard,R., Gay,C., Holt  
 ,I.E., Seede,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and  
 Quackenbush,J.

TITLE Assessment of gene expression patterns in a model of colon tumor  
 metastasis using a 19,200 element cDNA microarray

JOURNAL Unpublished (2000)  
 COMMENT Contact: John Quackenbush  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 3528  
 Fax: 301 838 0208  
 Email: johnq@tigr.org  
 Plate: 391

Seq primer: Forward.

Location/Qualifiers

FEATURES  
 source  
 1..699  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="MAGE resequences, MAGP"  
 /notes="Vector: pBluescriptSKm"

BASE COUNT 185 a 174 c 163 g 177 t  
 ORIGIN  
 Query Match 62.4%; Score 25.6; DB 10; Length 699;  
 Best Local Similarity 77.5%; Pred. No. 40;  
 Matches 31; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Y 2 CAGTCTCAGAGCGCTGAAATACCGAGCTGCTGAGC 41  
 378 CAGCTACTCAGAGCGCTGAGTGGGAGGACTGCTTGACC 417

RESULT 8  
 AA806631 453 bp mRNA linear EST 19-FEB-1998  
 DEFINITION ob71h08.s1 NCI CGAP GCBI Homo sapiens cDNA clone IMAGE:1336863 3'  
 similar to gb:K02067 H.sapiens mRNA for 781 RNA pseudogene (HUMAN  
 ); contains Alu repetitive element; contains element PTE5 repetitive  
 element ;, mRNA sequence.

ACCESSION AA806631  
 VERSION AA806631.1 GI:2875381  
 KEYWORDS EST.

SOURCE

human.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 453)

AUTHORS  
 TITLE

JOURNAL  
 COMMENT

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgaps-remail.nih.gov  
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,  
 Ph.D., Gerald Marti, M.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
 Bonaldo, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/HLNL at:  
 www-bio.llnl.gov/bbrp/image/image.html

Insert length: 758 Std Error: 0.00  
 Seq primer: -40ml3 fwd. BT from Amersham

High quality sequence stop: 363.

Location/Qualifiers

FEATURES

source

1..453  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1336863"  
 /clone\_lib="NCI CGAP GCBI"  
 /tissue\_type="germinal center B cell"  
 /lab\_host="DH103"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified  
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
 was prepared from human tonsillar cells enriched for  
 germinal center B cells by flow sorting (CD20+, IgD-),  
 provided by Dr. Louis M. Staudt (NCI), Dr. David Allman  
 (NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was  
 primed with a Not I - oligo(dT) primer  
 15'-TGTACCAATCTGAAGTGGAGCGCGCTCATTTTTTTTTTTT-3'  
 1. Double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified pT73 vector. Library  
 went through one round of normalization, and was  
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 105 a 107 c 116 g 125 t

ORIGIN  
 Query Match 61.5%; Score 25.2; DB 9; Length 453;  
 Best Local Similarity 78.9%; Pred. No. 48;  
 Matches 30; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 2 CAGTCTCAGAGCGCTGAATACCGAGCTGCTGAG 39

386 CAGCTACTCAGAGCGCTGAAGTGGGAGGACTGCTTGAG 423

Db

RESULT 9  
 AW900757/c 326 bp mRNA linear EST 24-MAY-2000  
 LOCUS CM1-NN1006-130400-195-cl2 NN1006 Homo sapiens cDNA, mRNA sequence.  
 DEFINITION  
 ACCESSION AW900757  
 VERSION AW900757.1 GI:8064962

KEYWORDS EST.

SOURCE

human.

ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 326)

AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,  
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
 ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
 Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

## COMMENT

Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the PAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-CM1-NN1006-130  
 400-185-cl2&t3=2000-04-13&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 12  
 High quality sequence stop: 326.  
 Location/Qualifiers

## FEATURES

source

1..326  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="NN1006"  
 /dev\_stage="Adult"

/note="Organ: nervous normal; Vector: puc18; Site 1: SmaI;  
 Site 2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the pUC 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

62 a 96 c 74 g 94 t

BASE COUNT

ORIGIN

Query Match 61.0%; Score 25; DB 10; Length 326;  
 Best Local Similarity 75.6%; Pred. No. 51;  
 Matches 31; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

1 TCAGTTCCTGAGAGCCTGAATACACGAGGACTCCCTGAGAC 41

|||||  
 213 TCAGTCTCTCAGAGGCTGAAGCAGGAGGATTGCTTGAGCC 173

RESULT 10

W070703/c

LOCUS AW070703 457 bp mRNA linear EST 13-OCT-1999  
 x229206.x1 NCI CGAP Br18 Homo sapiens cDNA clone IMAGE:2568250.3'  
 similar to contains Alu repetitive element; contains element HGR  
 repetitive element ; mRNA sequence.

ACCESSION

AW070703

AW070703.1 GI:6025701

VERSION

EST.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 457)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Chris Moskaluk, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life  
 Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The  
 I.M.A.G.E. Consortium DNA Sequencing by: Washington University  
 Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 www-bio.llnl.gov/bbrp/image/image.html  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 419.  
 Location/Qualifiers

## FEATURES

source

1..457  
 /organism="Homo sapiens"

/db\_xref="taxon:9606"  
 /clone="IMAGE:2568250"  
 /clone\_lib="NCI CGAP Br18"  
 /tissue\_type="four pooled high-grade tumors, including two  
 primary tumors and two metastatic to ovary"  
 /lab\_host="DH10B"  
 /note="Organ: breast; Vector: pCMV-SPORT6; Site 1: SmaI;  
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
 Library constructed by Life Technologies."

BASE COUNT

ORIGIN

Query Match 61.0%; Score 25; DB 10; Length 457;  
 Best Local Similarity 75.6%; Pred. No. 57;  
 Matches 31; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 TCAGTTCCTGAGAGCCTGAATACACGAGGACTCCCTGAGAC 41

|||||  
 Db 290 TCAGTCTCTCAGAGGCTGAAGCAGGAGGATTGCTTGAGCC 250

RESULT 11

AL575170

LOCUS

DEFINITION AL575170 LTI\_NFL006.PL2 Homo sapiens cDNA clone CS0DI063YB16.3  
 prime, mRNA sequence.

ACCESSION

AL575170

VERSION

AL575170.1

GI:12936074

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 836)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1..836

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="CS0DI063YB16"

/clone\_lib="LTI\_NFL006.PL2"

/tissue\_type="placenta"

/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA

was primed with a NotI-oligo(dT)-primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed by

Life Technologies. Contact : Feng Liang Life Technologies,

a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA Fax : (1) 301 610 8371

Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com"

BASE COUNT

ORIGIN

230 a 160 c 214 g 217 t

15 others

Query Match

60.0%; Score 24.6; DB 9; Length 836;

Best Local Similarity

73.2%; Pred. No. 98;

Matches 30; Conservative

1; Mismatches 10; Indels

0; Gaps

0;

QY 1 TCAGTTCCTGAGAGCCTGAATACACGAGGACTCCCTGAGAC 41

|||||

Db 587 TCAGTCTCTCAGAGGCTGAAGCAGGAGGATTGCTTGAGCC 627

RESULT 12

BF935457

LOCUS

BF935457

315 bp mRNA linear

EST 22-JAN-2001

```

DEFINITION CMI-NT0270-271200-678-e04 NT0270 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF935457
VERSION BF935457.1 GI:12352768
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 315)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/getatm2.pl?tl=CM1&t2=CM1-NT0270-
271200-678-e04&t3=2000-12-27&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 15
High quality sequence stop: 315.
FEATURES
source
location/Qualifiers
1..315
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NT0270"
/dev_stage="Adult"
/notes="Organ: nervous tumor; Vector: puc18; Site_1: Smal;
Site_2: Smal; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 78 a 70 c 88 g 79 t
ORIGIN
Query Match 58.5%; Score 24; DB 12; Length 315;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

1
2 CAGTGTCTGAGAGCTGTAATACAGAGCTGCTTGAGAC 41
|||||
246 CAGCTACTGGAGAGCTGACATGGTAGGACTGCTTGAGCC 285

RESULT 13
LOCUS AA369692 337 bp mRNA linear EST 21-APR-1997
DEFINITION EST01420 Prostate gland 1 Homo sapiens cDNA 3' end similar to EST
containing Alu repeat, mRNA sequence.
ACCESSION AA369692
VERSION AA369692.1 GI:2022225
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 337)
AUTHORS Adams,M.D., Soares,M.B., Kerlavage,A.R., Fields,C. and Venter,J.C.

```

```

TITLE Rapid cDNA sequencing (expressed sequence tags) from a
directionally cloned human infant brain cDNA library
JOURNAL Nat. Genet. 4, 373-380 (1993)
MEDLINE 94004965
COMMENT Other ESTs: BST81421 THC100858
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlavetigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (http://www.tigr.org/tdb/hgi/hgi.html)
Seq primer: M13-21.
FEATURES
source
location/Qualifiers
1..337
/organism="Homo sapiens"
/db_xref="ATCC (inhost):174463"
/db_xref="taxon:9606"
/clone_lib="Prostate gland 1"
/sex="male"
/dev_stage="adult, 21 yrs"
/notes="Organ: prostate; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI"
BASE COUNT 76 a 79 c 97 g 84 t 1 others
ORIGIN
Query Match 58.5%; Score 24; DB 9; Length 337;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

1
2 CAGTGTCTGAGAGCTGTAATACAGAGCTGCTTGAGAC 41
|||||
94 CAGCTACTGGAGAGCTGACATGGTAGGACTGCTTGAGCC 133

RESULT 14
LOCUS AW386852 374 bp mRNA linear EST 04-FEB-2000
DEFINITION CM0-PT0048-291299-141-c11 PT0048 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW386852
VERSION AW386852.1 GI:6891511
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 374)
AUTHORS HCGP http://www.ludwig.org.br/ORESTES.
TITLE The FAPESP/LICR Human Cancer Genome Project
JOURNAL Unpublished (1999)
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/getatm2.pl?tl=CM0&t2=CM0-PT0048-
291299-141-c11&t3=1999-12-29&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 11
High quality sequence stop: 374.
FEATURES
source
location/Qualifiers
1..374
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="PT0048"

```

